Department of Ophthalmology,
Odense University Hospital

Research Unit of Ophthalmology,
Department of Clinical Research,
Faculty of Health Sciences,
University of Southern Denmark

PhD Thesis

Sebastian Dinesen

Proliferative Diabetic Retinopathy: A Comprehensive Investigation Using Deep Learning, Epidemiology, and Retinal Imaging

Marts 2025

”Lediggang som saadan er ingenlunde en Rod til Ondt, tværtimod den er et sandt guddommeligt Liv, naar man ikke keder sig”.
- Søren Kirkegaard, 1853

Preface

Søren Kirkegaard havde måske fat i noget, da han i *Enten Eller* (1843) skrev, at livet ikke alene er produktion og travlhed, men fordybelse og eftertænksomhed. Tak til Amalie og mine sønner, Viggo og Eddie, for altid at minde mig om dette.

Contents

1 Aknowledgements

I want to express my sincere gratitude to my main supervisor, Jakob Grauslund, for his invaluable support and guidance throughout this PhD journey. Your ability to balance constructive feedback with encouragement has been instrumental in shaping both the project and my growth as a researcher. I am thankful for the opportunity you gave me to undertake this PhD.

I would like to extend my thanks to Tunde Peto for her support and commitment to my project. Your dedication was evident in the monthly check-ins, where you consistently offered thoughtful advice and guidance. Tunde has been an invaluable supervisor, not just for her expertise but also for her ability to see the person behind the project. A quality for which I am deeply grateful. Over time, I came to truly appreciate the depth of your knowledge and generosity in sharing it, which became increasingly vital as the project progressed and I learned how to make the most of your insights.

A sincere thank you goes to Thiusius Rajeeth Savarimuthu for opening the doors to his department, which enabled us to develop artificial intelligence (AI) algorithms under his leadership as head of the SDU Robotics. My deep gratitude goes to Jakob Andersen whose exceptional contributions in coding the algorithms were indispensable to the success of this project, and also to Andrea Blasi Núñez for her coding expertise and continuous support.

Lonny, Conny or something third. A loved child has many names. You are officially forgiven for calling me Simon in the very first days of my PhD project. I hope you can forgive me for calling you different names in return. Without you the project would have never succeeded, so I cannot thank you enough for your commitment to me and the project.

To Yousif, thank you for your constant positivity and sharp inputs throughout my PhD project. This is truly appreciated.

To my fellow PhD student and undergraduate research students along the way, thank you the discussions, laughs and the much-needed coffee breaks. I feel grateful to say that I gained true friendships during my time as a PhD student. Anne, you are truly a wonderful person, and I am really happy to see you have

found your path. Thank you for the time and talks we had as office buddies. To Simon, it has been a pleasure and very easy to share office with you. I am looking forward to you getting me in gold standard shape over the summer. To Marianne, even though we shared an office for less than a year, it did not take long to appreciate your delightful character. Hold on to that. And the boys, Frederik and Benjamin, it has been a pleasure traveling the world with you, but I also want to thank you for the great discussions we had over the years. I hope to see you around.

2 Publications

Paper I

Five-year Incidence of Proliferative Diabetic Retinopathy and Associated Risk Factors in a Nationwide Cohort of 201,945 Danish Patients with Diabetes
Ophthalmology Science 2023 Feb 24; 3(3):100291.
Sebastian Dinesen**,** Lonny Stokholm, Yousif Subhi, Tunde Peto, Thiusius Rajeeth Savarimuthu, Nis Andersen, Jens Andresen, Toke Bek, Javad Hajari, Steffen Heegaard, Kurt Højlund, Caroline Schmidt Laugesen, Ryo Kawasaki, Sören Möller, Katja Schielke, Anne Suhr Thykjær, Frederik Pedersen, Jakob Grauslund.

Paper II

Retinal main vessel calibers and systemic markers for long-term development of proliferative diabetic retinopathy
Acta Ophthalmologica 2023;00:1-7
Sebastian Dinesen, Lonny Stokholm, Yousif Subhi, Jan Erik Henriksen, Thiusius Rajeeth Savarimuthu, Tunde Peto, Jakob Grauslund.

Paper III

Prediction of Proliferative Diabetic Retinopathy from Retinal Images using Deep Learning
Sebastian Dinesen, Lonny Stokholm, Andrea Blasi Núñez, Jakob Kristian Holm Andersen, Kristian Nielsen, Thiusius Rajeeth Savarimuthu, Tunde Peto, Jakob Grauslund.
(Included as a manuscript)

Paper IV

A Deep Learning Segmentation Model for Detection of Active Proliferative Diabetic Retinopathy
Sebastian Dinesen, Marianne G. Schou, Christoffer V. Hedegaard, Yousif Subhi, Thiusius R. Savarimuthu, PhD, Tunde Peto, Jakob K. H. Andersen, Jakob Grauslund

3 Abbreviations

AI	Artificial intelligence
ANN	Artificial neural networks
AUC	Area under the curve
ATC	Anatomical Therapeutic Chemical
CI	Confidence Interval
CNN	Convolutional neural networks
CPR number	Civil personal registration number
CRAE	The central retinal arteriolar equivalent
CRVE	The central retinal venular equivalent
CVAT	Computer Vision Annotation Tool
DiaBase	Danish Registry of Diabetic Retinopathy
DL	Deep learning
DR	Diabetic retinopathy
FDDB	Funen Diabetes Database
HR	Hazard ratio
ICDR	International Clinical Diabetic Retinopathy
ICD-10	International Classification System of Disease
IRMA	Intraretinal microvascular abnormalities
NPV	Negative predictive value
OR	Odds ratio
PDR	Proliferative diabetic retinopathy
PPV	Positive predictive value
VEGF	Vascular endothelial growth factor

4 Summary

Over 828 million people are diagnosed with diabetes worldwide. Diabetic retinopathy (DR) is a microvascular complication of diabetes that develops in many patients after prolonged exposure to hyperglycaemia. DR progresses through five stages starting with no DR (level 0), minimal DR (level 1), moderate non-proliferative DR (NPDR) (level 2), severe NPDR (level 3) and the potential sight-threatening end-stage proliferative diabetic retinopathy (PDR) (level 4).

The dysfunction of the microvasculature arises from various pathological mechanisms that leads to retinal ischemia. This causes the formation of fragile new vessels prone to bleeding into the vitreous body, which can damage the vision. It is important to diagnose and treat these new fragile vessels before irreversible vision loss occurs. Patients are therefore invited for eye screening at appropriate intervals based on their DR level and HbA1c.

This Ph.D. thesis aimed to investigate PDR from multiple perspectives, including an epidemiological approach and the development of diagnostic methods for identifying and predicting PDR.

We performed an epidemiological analysis of a nation-wide population of 201,945 patients in Paper I, which included estimates of prevalence and incidence rates as well as risk factor for progression to PDR. In Paper II and III, we aimed to predict future development of PDR from retinal images prior to the date of diagnosis. This included retinal imaging analysis with a semi-automatic software to analyse retinal vessel diameters as predictors for PDR as well as developing a deep learning (DL) model for the same purpose. In Paper IV, we investigated whether a DL segmentation model could detect active PDR by identifying new vessels and/or preretinal haemorrhages.

The first study identified key risk factors for PDR and found lower incidence rates compared to earlier periods. However, the second and third studies did not successfully predict PDR. The fourth study developed a high-performing DL segmentation model for active PDR detection.

The findings in this thesis provide valuable insights into PDR in a well-regulated and low-risk population. The lower incidence rates are promising for the future, but the growing prevalence of diabetes underscores the need for continuous screening efforts. The projected workload increase could be managed through the implementation of a DL model. The predictive potential for DL requires further exploration, particularly through the incorporation of clinical and demographic risk factors, as a clear opportunity to enhance accuracy.

DL is a promising method to implement to help manage the increase in workload projected with increasing diabetes prevalence. Further exploration of DL's predictive potential, incorporating additional risk factors, is essential for improving its accuracy in PDR prediction.

5 Motivation

Several aspects convinced me to commit myself to this PhD project. I was aware that during the process, there would be situations where I would find myself in deep waters, but the possibility to strengthen my professional and academic skills has been an essential driving force from the beginning.

The studies within this thesis and their potential for positive changes of the DR screening program have been a strong source of motivation. The burden of DR increases worldwide, and the need for screening is obvious, but the economic resources for medical care are unevenly distributed[1]. Hopefully, the results from this thesis can also provide insights for countries with limited medical resources. Whether the focus has been on identifying risk factors for PDR predicting future disease, detecting current disease, or freeing up healthcare resources to direct focus for patients most in need, the goal of enhancing the quality of the Danish screening program for DR, thereby preventing vision impairment and blindness, has remained a key driving force.

6 Background

6.01 Diabetes Mellitus

Diabetes is a chronic metabolic disorder characterized by impaired insulin production or function that disrupts normal blood glucose levels. Type 1 diabetes is an autoimmune disease that leads to destruction of the insulin producing beta cells in the pancreas. Type 2 diabetes is a complex condition defined by insulin resistance and/or decreased insulin production and accounts for more than 90% of cases. The estimated global prevalence of both type 1 and 2 diabetes was 9.3% (463 million patients) in year 2019 projected to increase to 10.2% (578 million) in 2030 and 10.9% (700 million) in 2045[2]. A more recent estimate from 2022 indicates that 828 million adults lives with diabetes, which suggest an even greater disease burden than previously reported[3].

6.02 Diabetic Retinopathy

DR is a common microvascular complication of diabetes that progresses through two main stages: a non-proliferative stage and a proliferative stage. The latter can be a serious threat to vision if not diagnosed and treated on a timely manner[4]. In 2012, the estimated global prevalence of any form of DR was 77.2% in patients with type 1 diabetes and 25.2% in those with type 2 diabetes, with a combined prevalence of 22.3% reported in a more recent study from 2021[5]. PDR was reported in 32.4% of patients with type 1 and 3.0% of those with type 2 diabetes[6]. The incidence of PDR has been declining in western countries over the last few decades thanks to screening, early diagnosis and better treatment options for diabetes per se, as shown in Figure 1[7].

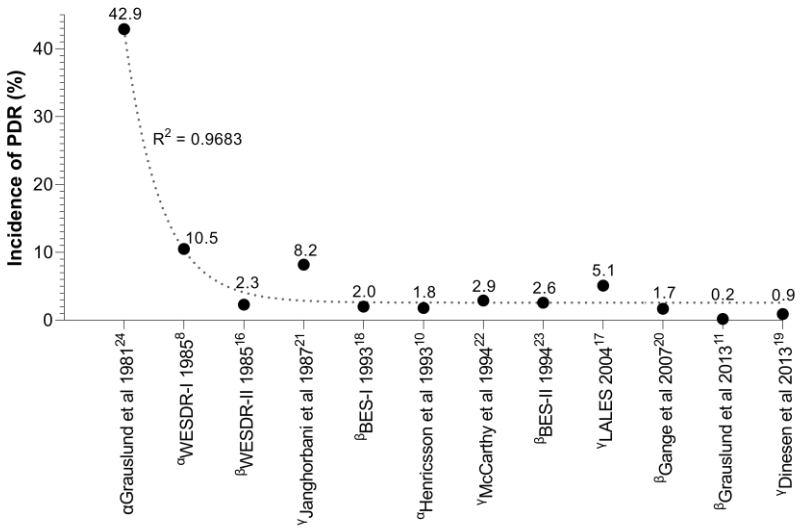


Figure 1: Trends in the incidence of proliferative diabetic retinopathy in population-based studies of type 1[α] or 2 diabetes[β] and some including both types of diabetes[γ]. The year marks the baseline date of each follow-up period and the uppercase number is the reference number. PDR = proliferative diabetic retinopathy. Reprinted from Abou Taha et al [3] with allowance as a co-author and with permission from BMC.

The Wisconsin Epidemiologic Study of Diabetic Retinopathy (WESDR) identified the association between hyperglycaemia and progression of DR in the late 1980's[8]. Since then, the monitoring and treatment of diabetes have improved, including long-term monitoring of blood glucose with glycosylated haemoglobin (HbA1c)[9], treatment regimens with basal and bolus-insulin, pumps with automated insulin infusion, patient education and implementation of screening programs[4,10]. These improvements in controlling high or unstable blood glucose, along with better regulation of other well-known risk factors for DR progression such as hypertension, hypercholesterolemia, obesity and smoking, reasonably explains the decline in the incidence of PDR in high income countries[6].

Chronic hyperglycaemia causes the development of PDR through multiple pathophysiological mechanisms, including vascular endothelial dysfunction, inflammation and neurodegeneration[4]. Damage to the endothelial cells lining the retinal vasculature compromises the inner blood-retina barrier, leading to vascular leakage and the formation of diabetic macular oedema, which can cause

severe vision loss. Additionally, capillary occlusion causes retinal ischemia, prompting the release of proangiogenic factors like vascular endothelial growth factor (VEGF) in affected areas. VEGF stimulates the formation of new blood vessels, typically on or around the optic disc or in the peripheral retina. New vessels are fragile and a potential threat to vision as they can cause vitreous haemorrhage, vitreous contraction and tractional retinal detachment[4,11].

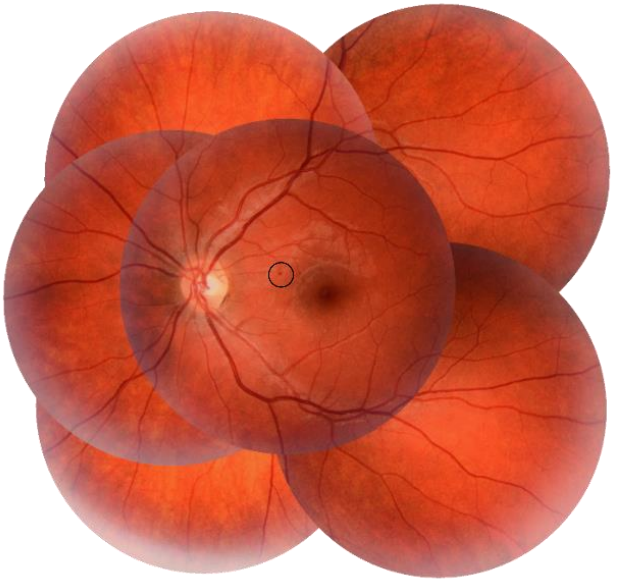
Laser photocoagulation revolutionized the treatment of PDR 50 years ago[12] by decreasing the ratio of proangiogenic to antiangiogenic factors, and remains standard of care today used alongside VEGF inhibition and/or vitrectomy[11].

6.03 Diabetic Retinopathy Screening

The purpose of this section is to introduce the Danish screening framework, as the studies included in this thesis were designed specifically for this context.

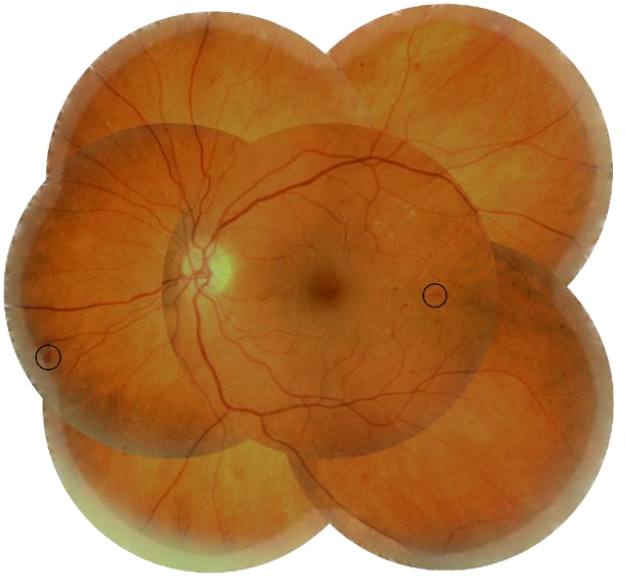
The World Health Organization lists several criteria that must be fulfilled before recommending screening for a given disease[13]. DR meets all these criteria for screening, and many countries have implemented DR screening programs. The Danish program is a national solution, free of charge for the patients and based on national clinical guidelines[14,15]. The purpose is to identify the pathological changes that are threat to sight and to treat these before they result in irreversible vision loss.

The Danish screening program conducts more than 100,000 screening episodes every year. It offers lifelong attendance and the clinical examinations are carried out by both practicing ophthalmologists and hospital departments. DR is classified according to the International Clinical Diabetic Retinopathy (ICDR) Severity Scale, which ranges from 0 to 4 with higher numbers indicating greater disease severity (Figure 2)[16]. I applied the ICDR Severity Scale in the studies forming this thesis.



ICDR 1

One or more microaneurysms

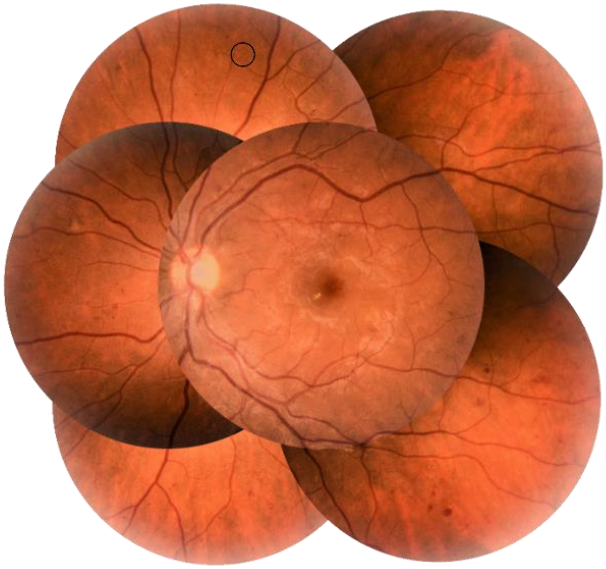


ICDR 2

More than microaneutysms, but less than ICDR level 3

ICDR 3

> 20 haemorrhages in each of 4 quadrants
or
venous beading ≥ 2 quadrants
or
IRMA ≥ 1 quadrant



ICDR 4

Neovascularization and/or pre-retinal bleeding

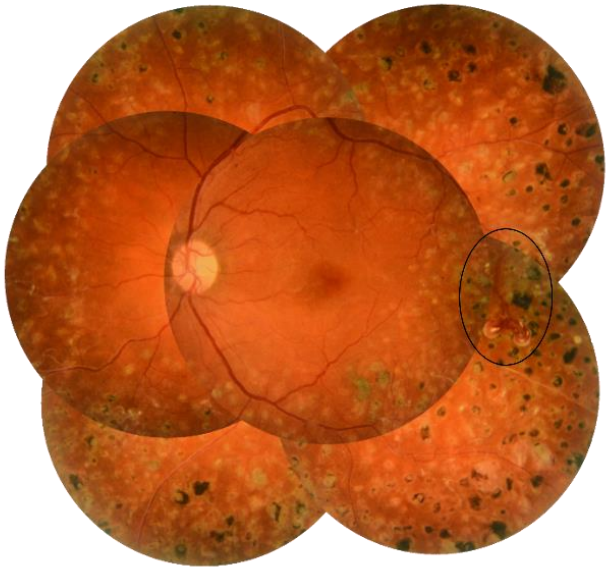


Figure 2: Six-field images of the retina displaying levels 1 to 4 on the International Clinical Diabetic Retinopathy (ICDR) Severity Scale. Level 0, which indicates no diabetic retinopathy, is not represented in the figure. Higher numbers on the scale indicates increasing severity. IRMA = Intraretinal microvascular abnormalities. Image origin: Steno Diabetes Centre Odense.

The requirements for a Danish screening episode are measure of visual acuity and a minimum of two 45-degree retinal images, but in hospital settings the standard of care is six-field images. Optical coherence tomography is performed if fundus photography raises suspicion of diabetic macular oedema. Although diabetic macular oedema is mentioned here and is a central part of the clinical assessment in the diabetic eye, it falls out of the scope of this thesis, and will not be explored further.

Clinicians adjust the flexible and individual interval for screening episodes based on the DR severity and the biomarker HbA1c. The screening intervals listed below are the current recommendations by the Danish national guidelines[14,15].

ICDR level 0 or 1

Screening every 24 months

ICDR level 2

Screening every 12 to 24 months

If HbA1c > 80mmol/mol: Screening every 6 to 12 months

ICDR level 3

Screening every 3 to12 months

If HbA1c > 80mmol/mol: Screening every 3 months

ICDR level 4 (Stable or treated PDR)

Screening every 6 to 12 months

If HbA1c > 80mmol/mol: Screening every 3 to 12 months

New or recurrent PDR

Direct referral to eye department at a hospital

6.04 Deep Learning

Machine learning technologies are present in many everyday situations in modern society, ranging from smartphones and websites to self-driving cars and generative language models like the ChatGPT. Traditional machine learning models were limited in their ability to extract features from raw data, but this changed with representation learning methods that allowed machines to be fed with raw data and automatically discover the features for classification[17]. DL is such a representation model based on artificial neural networks (ANN). This type of AI enables machines to perform automatic extraction of features and patterns from large datasets[17]. DL has revolutionized several domains with its applications. In computer vision, it performs object detection and classification of images[18], while it transcripts from speech recognition[19], and translates from natural language processing[20].

Artificial Neural Networks

ANNs mimic the structure and function of the human brain, being composed of layers of interconnected nodes (neurons). For each connection, a linked *weight* determines the strength of the signal passed from one neuron to the next. Three types of layers define an ANN, where the first layer, the input layer, is responsible for processing the raw data. The second set of layers is the hidden layers, and it is the depth of these (number of hidden layers) that defines a DL network, which is characterized by having two or more hidden layers. The final layer, the output layer, generates the relevant results based on the purpose of a given DL model[17].

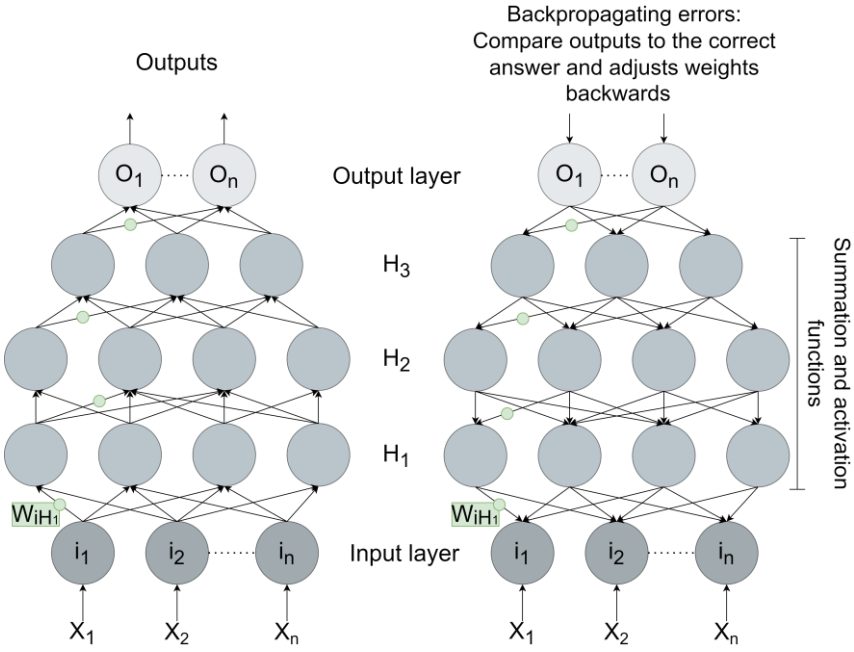


Figure 3: Structure of a deep learning model. A standard feedforward concept consisting of an input, multiple hidden layers and an output layer. The green dots represent the weights connected with each neuron. The right-sided diagram illustrates the backpropagation process used to correct errors by adjusting the weights in the network. Self-made figure with diagrams.net.

Learning Process

Forward propagation is the step where data flows through the different layers of the DL network (Figure 3). Each neuron receives inputs from the neurons in the previous layer to which it is connected. The strength of the signal transmitted from a single neuron is determined by the weight assigned to the connection, meaning that higher weight values have a greater impact on the output of the neuron. A summation function calculates the values a neuron receives from its connected neurons. This value passes through the activation function, which determines whether the neuron will transmit a signal. The activation function is non-linear, allowing a DL algorithm to learn and represent more complex and abstract patterns. As for example, a DL algorithm analysing a retinal image of PDR must be able to recognize a new vessel regardless of its position, orienta-

tion or illumination. This is possible by the non-linearity introduced by the activation function[17].

The *loss function* plays an important role in the learning process by helping the algorithm identify its errors. It generates a numerical value based on the DL algorithms performance. If the results is far from the true answer, it produces a high number indicating more error[17].

The *backpropagation* process sends the error information from the loss function backwards through the ANN to adjust the weights of the connections between neurons. This process is repeated many times to gradually improve the models performance[21].

Convolutional Neural Networks

Convolutional neural networks (CNN) are a type of DL model particularly designed for computer vision tasks, such as image processing. They are called "convolutional" because they rely on a convolutional mathematical operation as a part of its key architecture. This process involves sliding a filter (also known as a kernel) over the entire image to extract features from localized regions like patterns, edges, textures or corners one image patch at a time. This is what differentiates CNNs from other ANNs, namely that they analyse images in smaller parts (patches) rather than the whole image at once. This allows CNNs to recognize specific patterns, regardless of their location in the image, which is important if you want to achieve automated detection and segmentation of new vessels. The generated *feature maps* from the filter application highlight the important features in the input image, and the connected weights are shared across the entire image during the convolutional process. The weight sharing reduces the number of parameters the CNN needs to learn, making it very efficient[17].

The Black Box

The Black box phenomenon describes the lack of transparency in how DL models arrives at their predictions. The processes that transform input data into a final model prediction (outcome) are hidden in traditional black box models, which makes it difficult for humans to understand how it processes data and makes a decision[22].

6.05 Danish Health Registers

Denmark has a long tradition of register-based research, which was strongly enhanced in 1968 with the introduction of the Civil Personal Registration (CPR) system. This system assigns a unique personal identifier to every individual born in Denmark and enables data linkage across various registers (Figure 4). The CPR register contains data on sex, date of birth, civil status, migration and vital status.

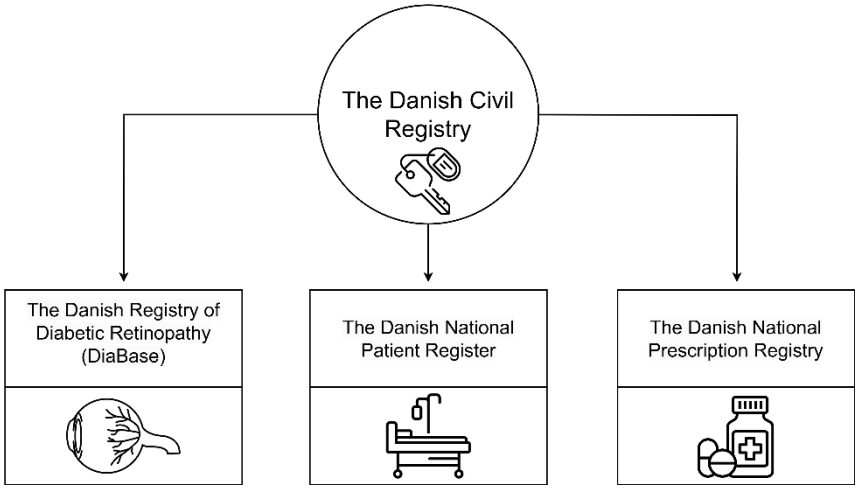


Figure 4: The Danish registers forming Paper I. The Danish Civil Registry was used to connect data across the three pictured registers at the individual level with the unique civil personal register number. The figure was selfmade with diagrams.net.

The Danish Registry of Diabetic Retinopathy (DiaBase) is a national database with information from individual clinical visits in the Danish screening program for DR since 2013. The data included includes DR severity level, date of screening episode and visual acuity[23]. The DiaBase is reliable on the clinician's effort to report the results from a given screening episode and receives information from both practicing ophthalmologists and hospitals.

The Danish National Patient Registry was established in 1976 and includes individual data on diagnosis, surgery and treatments from inpatients, outpatients as well as emergency patients treated at a Danish hospital. Since year 1994, the International Classification System of Disease (ICD) has been used as coding system[24]. The register contains a primary (A) diagnosis or treatment code, secondary (B) diagnosis or treatment codes in connection with the A-code[24].

The Danish National Prescription Registry is a national database that contains individual-level information on all prescription drugs dispensed in pharmacies. The dispensed drugs are categorized by the Anatomical Therapeutic Chemical classification (ATC) code, which enables specific drug identification[25].

The Funen Diabetes Database (FDDB) was established in 2003 and covers 80% of the patients diagnosed with diabetes in Funen area[26]. It includes data from routine clinical diabetes examinations including the CPR identifier, date of contact, age, and sex. Clinical data encompass DR levels (ICDR levels 0 to 4), foot examination results, blood pressure, weight, height, body mass index, and smoking status. Biomarkers are automatically transferred from laboratory systems to the FDDB and include HbA1c, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, total cholesterol, triglycerides, plasma creatinine, urine creatinine, urine albumin concentration and the urine albumin to creatinine ratio. Additionally, the database records information on prescribed insulin and oral glucose-lowering medications, as well as diabetes-related complications, including cardiovascular disease, diabetic nephropathy, hypoglycaemia, and ketoacidosis.

7 Purpose

The overall purpose of this thesis was to enhance our understanding of PDR in a nation-wide diabetes population. This included an epidemiological approach to explore risk factors, prevalence, incidence rates and developing innovative diagnostic methods for predicting and detecting PDR.

Table 1: Overview of the studies included in the PhD thesis

	Paper I	**Paper II**
Design	Register-based cohort study	Retrospective case-control study
Population	201,945 patients	128 patients
Sample size	401,462 eyes	128 retinal images
Subjects	Patients with type 1 or 2 diabetes	Patients with type 1 or 2 diabetes Cases: Subsequent PDR Controls: No PDR
Inclusion period	2nd of January 2013 to 30th of December 2018.	2nd of January 2003 to 22nd of November 2019.
Inclusion criteria	Type 1 or 2 diabetes, aged ≥ 18 years of age and index date between 2nd January 2013 to 30th December 2018	Type 1 or 2 diabetes, ICDR-level 0 or 1 at baseline and subsequent progression to PDR (cases)
Data sources	DiaBase, The Danish National Patient Register, The Danish National Prescription Registry, The Danish Civil Registration System	The Funen Diabetes Database, Steno Diabetes Centre Odense image database
Primary outcome measures	Hazard ratios of markers of incident PDR and 1, 3 and 5 year PDR incidence rates	Odds ratio for progression to PDR
Statistics	Uni- and multivariable Cox proportional hazard models & Temporal analysis of the 1, 2 and 3 year risk of progression to PDR	Uni- and multivariable logistic regression models & Intergrader agreement: Two-way mixed-effects model

PDR = proliferative diabetic retinopathy, ICDR scale = International Clinical Diabetic Retinopathy Severity Scale

Paper III	Paper IV
Deep learning diagnostic accuracy study	Deep Learning diagnostic accuracy study
960 patients	199 patients
26,828 retinal images	938 retinal images
Patients with type 1 or 2 diabetes Cases: Subsequent PDR Controls: No PDR	Patients with type 1 or 2 diabetes
2nd of January 2013 to 30th of May 2022	2nd of January 2003 to 22nd of November 2019
Type 1 or 2 diabetes.	Type 1 or 2 diabetes, inactive (controls) or active (active) PDR
DiaBase, Steno Diabetes Centre Odense image database	The Funen Diabetes Database, Steno Diabetes Centre Odense image database
Sensitivity and specificity for prediction of PDR	Sensitivity and specificity for detection of active PDR
Sensitivity and specificity calculated as measures of diagnostic accuracy	Sensitivity and specificity calculated as measures of diagnostic accuracy

8 Paper I

8.01 Objective

To evaluate the incidence of PDR and identify risk factors associated with progression to PDR in a national population of Danish patients with diabetes.

8.02 Methods

Study Design

This national five-year cohort study examined eyes from patients with and without PDR progression, using data Danish registers, including the CPR system, DiaBase, the National Patient and the National Prescription Registry.

Study Population and Data Sources

The study population consisted of patients aged 18 years or older with type 1 or 2 diabetes who participated in the Danish DR screening program and had their visits recorded in the DiaBase from the 2th of January 2013 to the 31th of December 2018.

Inclusion and Exclusion Criteria

Patients were followed from their index date (first screening) until their last screening or an endpoint: death, PDR progression, or emigration. Individuals contributed one or both eyes, but were represented by their worst eye in Tables 2 and A1. Eyes with PDR at baseline, regression from PDR (if an individual is reported with DR level 1 to 3 after diagnosed with PDR) or missing DR level assessments at any screenings were excluded.

Statistical Analysis

Table 2 shows the comparison of baseline characteristics for cases (PDR) versus controls (ICDR 0 to 3). Continuous variables are presented as medians with interquartile ranges and categorical variables as counts with proportions.

We used a multivariable Cox proportional hazards model to evaluate associations between baseline variables and progression to PDR, adjusting for sex, age, civil status, diabetes type, anti-hypertensive and cholesterol lowering medications as well as a modified Charlson Comorbidity Index score excluding diabetes. Given that most patients (98.8%) contributed two dependent eyes, we applied robust standard errors to account for the clustered structure.

We calculated the one, three and five-year incidences for progression to PDR from the baseline DR level in a temporal analysis (Table 3).

A competing risk analysis was performed as we were suspicious of the low progression rates from DR level 3. A Fine-Gray sub distribution hazard model was applied to account for mortality as a competing risk in the progression to PDR among patients with DR level 3, compared to those with DR level 2.

8.03 Results

The five-year cohort included 201,945 patients of whom 1,780 (0.9%) patients and 2,384 (0.6%) eyes had progression to PDR. The rate of PDR progression varied by baseline DR-level: 0.1% for level 0 (510 eyes), 2.7% for DR-level 1 (1,041 eyes), 4.7% for DR-level 2 (535 eyes) and 14.7% for DR-level 3 (298 eyes). The median number of visits per patient was three (interquartile range 1-4).

The baseline characteristics categorized by DR levels can be studied in Table A1 (appendix). Table 2 presents the baseline differences between patients with and without subsequent progression to PDR. Male sex (60.3 vs. 56.5%, p = <0.001), lower age (59.6 vs. 65.8 years, p = <0.001), longer duration of diabetes (type 1 diabetes: 19.9 vs. 14.9 years, p = <0.001; type 2 diabetes: 11.7 vs. 5.3 years, p = <0.001) and type 1 diabetes (28.7% vs. 7.6%, p = <0.001) were all associated with PDR progression. Individuals who were never married (21.9 vs. 14.8%, p = <0.001), used insulin (86.3 vs. 31.5%, p = <0.001) and were prescribed anti-hypertensive medicine (80.0 vs. 74.6%, p = <0.001) were also in

increased risk for subsequent PDR incidence.

Table 2: Baseline Differences between Patients with and without Subsequent Progression to PDR

	All	Progression to PDR	No progression to PDR	P-value
Number of patients, n	201,945	1,780	200,165	
Male sex, n (%)	114,107 (56.5)	1,074 (60.3)	113,033 (56.5)	0.001
Age (y), median (IQR)	65.8 (55.6;73.2)	59.6 (48.0;68.9)	65.8 (55.6;73.2)	0.001
Diabetes type, n (%)				<0.001
Type 1 diabetes	15,728 (7.8)	511 (28.7)	15,217 (7.6)	
Type 2 diabetes	152,622 (75.6)	421 (23.7)	152,201 (76.0)	
Unknown	33,595 (16.6)	848 (47.6)	32,747(16.4)	
Duration of diabetes, years (IQR)				<0.001
Type 1 diabetes	15.4 (6.7;20.2)	19.9 (19.1;21.3)	14.9 (6.5;20.2)	
Type 2 diabetes	5.3 (2.0;9.7)	11.7 (5.8;17.4)	5.3 (2.0;9.7)	
Marital status, n (%)				<0.001
Never married	30,042 (14.9)	390 (21.9)	29,652 (14.8)	
Married	116,592 (57.7)	1,005 (56.5)	115,587 (57.7)	
Widowed or divorced	55,311 (27.4)	385 (21.6)	54,926 (27.4)	
Charlson Comorbidity Index score, n (%)				<0.001
0 (low)	147,310 (72.9)	770 (43.3)	146,540 (73.2)	
1 (moderate low)	25,987 (12.9)	698 (39.2)	25,289 (12.6)	
2 (Moderate high)	18,153 (9.0)	163 (9.2)	17,990 (9.0)	
≥ 3 (high)	10,495 (5.2)	149 (8.4)	10,346 (5.2)	
Use of medication, n (%)				
Insulin	64,672 (32.0)	1,536 (86.3)	63,136 (31.5)	<0.001
Glucose-lowering	154,069 (76.3)	759 (42.6)	153,310 (76.6)	<0.001
Anti-hypertensiva	150,825 (74.7)	1,424 (80.0)	149,401 (74.6)	<0.001
Cholesterol lowering	149,219 (73.9)	1,294 (72.7)	147,925 (73.9)	0.29

IQR = Interquartile range. PDR = proliferative diabetic retinopathy. DR = diabetic retinopathy. Adapted and reprinted from Paper I.

In a Cox proportional hazards model comparing baseline characteristics between patients with and without progression to PDR, we found that those who progressed had a longer duration of diabetes (hazard ratio (HR) 4.66 per 10 years; 95% confidence interval (CI), 4.05 to 4.37), Charlson Comorbidity Index score over 0 (score 1: HR 4.62, 95% CI, 4.14 to 5.15; score 2: HR 2.28; 95% CI, 1.90 to 2.74; score ≥3: HR 4.28; 95% CI, 3.54 to 5.17) were more likely to be unmarried (HR 1.36; 95% CI, 1.19 to 1.56) and to have type 1 diabetes (HR 9.61; 95% CI, 8.01 to 11.53). In addition, patients who received insulin (HR 5.33; 95% CI, 4.49 to 6.33) or anti-hypertensive (HR 2.23; 95% CI, 1.90 to 2.61) treatment showed an increased risk of progression to PDR.

Patients with lower age (HR 0.87; 95% CI, 0.83 to 0.91) at their first screening visit and those medicated with non-insulin glucose lowering drugs (HR 0.51; 95% CI, 0.45 to 0.57) had a decreased risk of PDR progression. Sex and cholesterol lowering medicine showed no difference in risk of progression.

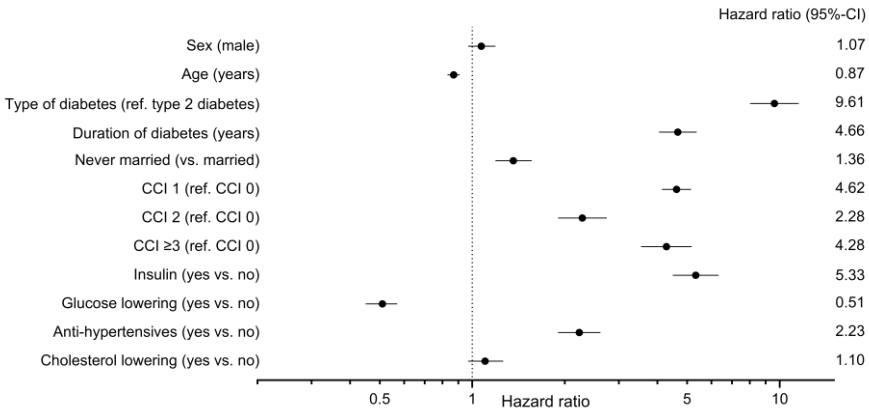


Figure 5: Forrest plot of hazard ratios for progression to proliferative diabetic retinopathy by baseline characteristics. Shown are a multivariable Cox proportional hazard model. Cholesterol lowering includes all cholesterol lowering medications. Glucose lowering is excluding insulins. CCI = Charlson Comorbidity Index. Ref. = reference. Adapted and modified from paper I.

The temporal analysis demonstrated that the risk of progression to PDR increased with higher baseline DR-levels and longer follow-up periods. Over a one-

year period, eyes with baseline DR-level 0 had a low incidence rate of 0.6 per 1000 person-years compared to incidence rates of 8.3 per 1000 person-years for DR level 1, 13.9 per 1000 person-years for DR-level 2., whereas those with DR-level 3 had a markedly higher incidence rate of 52.7 per 1000 person-years. This trend persisted over three- and five-year periods, as shown in Table 3.

Table 3: Incidence Rates of PDR Progression at One, Three and Five Years

One year				
Level of DR	Eyes at risk	Events	Observations time	Incidence rate per 1000 person years
All	401,462	570	282,631	2.0
0	352,143	138	243,885	0.6
1	36,280	239	28,800	8.3
2	11,006	119	8,542	13.9
3	2,033	74	1,403	52.7
Three years				
Level of DR	Eyes at risk	Events	Observations time	Incidence rate per 1000 person years
All	401,462	1,759	680,700	2.6
0	352,143	382	584,546	0.6
1	36,280	778	71,903	10.8
2	11,006	377	20,979	17.9
3	2,033	222	3,270	67.9
Five years				
Level of DR	Eyes at risk	Events	Observations time	Incidence rate per 1000 person years
All	401,462	2,384	817,743	2.9
0	352,143	510	696,917	0.7
1	36,280	1,041	89,823	11.6
2	11,006	535	26,900	19.9
3	2,033	298	4,101	72.7

PDR = proliferative diabetic retinopathy. Adapted and modified from Paper I

The risk of mortality before progression to PDR was considered an alternative event for those with baseline DR-level 3. Therefore, we compared the risk of progression between DR-level 2 and 3 showing higher risk for progression with DR-level 3 at baseline (HR 3.25; 95% CI, 2.74–3.87). The HR remained unaffected when we tested death as a competing event to PDR progression (HR 3.25; 95% CI, 2.74–3.87).

9 Paper II to III

9.01 Objective

The aim of the two studies was to explore whether PDR could be predicted from retinal images using semi-automatic geometrical vessel analysis (Paper II) and DL (Paper III).

Paper II aimed to predict the development of PDR from early-stage DR images (ICDR levels 0 to 1) while **Paper III** aimed to predict PDR from images across all DR stages (ICDR levels 0 to 3) and from subsets of ICDR levels 1, 2, and 3.

9.02 Methods

Study Designs

Paper II used a retrospective matched case-control design while **Paper III** was a DL-based diagnostic accuracy study.

Study Populations and Data Source

Paper II: The FDDB supplied data from 2003 to 2019 forming the study populations while the retinal images originated from our local DR screening image database at Odense University Hospital. It also utilized clinical and demographic data from the FDDB as detailed in Table A2.

Paper III: The population consisted of patients with type 1 or 2 diabetes and was established using data from the DiaBase[23]. We defined the PDR endpoint in DiaBase by identifying eyes classified with DR level 4 and also included date of PDR diagnosis, date of first screening episode and baseline DR level. Eyes from those without a PDR diagnosis were included as controls.

Inclusion and Exclusion Criteria

Both studies required cases to have PDR and we only included images from screening episodes prior to the date of PDR diagnosis.

In Paper II, we first identified cases as individuals registered in the FDDB with PDR (DR level 4) and a baseline DR level of 0 or 1. Individuals could contribute one or both eyes. We conducted retrospective manual searches in our local

image database to retrieve the initial retinal image, ensuring date alignment between the image database and the FDDB (Figure 6). It led to exclusion of patients when the initial retinal image was unavailable. The first screening date served as baseline date for our case population that we now matched 1:3 to controls. The control patients had the same baseline DR-level, sex, age (± five years), type of diabetes and observation time (± one year) in the FDDB.

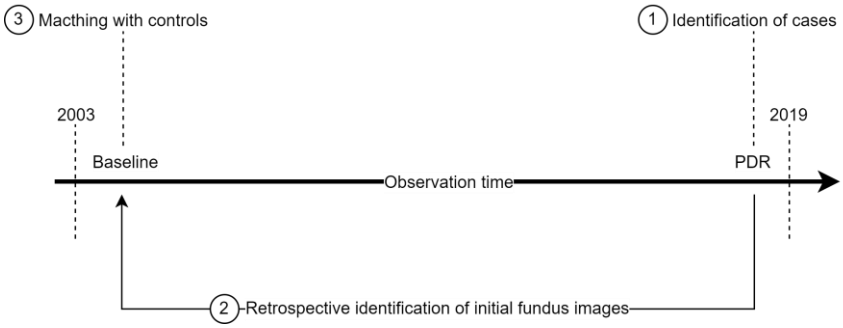


Figure 6: A timeline displaying the formation of the study population. 1: Identification of patients with proliferative diabetic retinopathy in the Funen Diabetes Database from year 2003 to 2019. 2: Retrospective identification of initial fundus images. 3: Matching of cases to controls with the first screening episode as baseline. Adapted and modified from paper II.

For **Paper III**, we included retinal images from our localized database at Odense University Hospital (OUH). The retinal images are labelled with CPR-number and date of screening. When creating the PDR image dataset, we only included images preceding the date of diagnosis and excluded those beyond the date of diagnosis. An individual could contribute several images, but from different screening dates, and could be included with both eyes. Images from eyes without a PDR diagnosis were included as controls.

Retinal Image Preparation

A certified grader (SD) manually confirmed the PDR diagnosis in both studies.

Paper II included an image quality assessment and an intergrader agreement analysis. We evaluated the binary black-and-white vessel map to determine the VAMPIRE software's ability to accurately track vessels in the retinal images. No images were excluded. Further, a VAMPIRE intergrader analysis of 38 (25%) random images displayed high agreement of 90% for CRAE and 94% for CRVE.

In **Paper III**, we split the images into training (70%), validation (20%), and testing (10%) datasets for model development. In addition to developing a model using all DR levels combined, separate PDR prediction models were also developed for ICDR levels 1, 2, and 3 (Table 4). An individual was allowed to contribute more than one image, but could only represent one of the three datasets.

Both studies included high-resolution (6,528 x 6,528 pixels) six-field retinal images.

Retinal Image Analysis

This thesis applied two distinct methods to analyse retinal images as predictors of progression to PDR.

In Paper II, we used the Vessel Assessment and Measurement Platform for Images of the Retina (VAMPIRE; version 3.2; The VAMPIRE Group; Edinburgh, United Kingdom). The software can automatically identify retinal vessels and anatomical landmarks such as the optic disc and macula. This facilitated the construction of a grid with circular zones around the optic disc. These zones were defined as follows: Zone A encompassed 0.0–0.5 disc diameters from the optic disc margin, Zone B covered 0.5–1.0 disc diameters, and Zone C extended 0.5–2.0 disc diameters. We manually adjusted occasional inaccuracies in the automated detection as necessary. Vessels were excluded from analysis under specific circumstances, including ambiguous vessel classification, instances where two vessels were incorrectly identified as a single vessel, or when a vessel from the choroidal layer was misclassified as a retinal vessel by the software. The central retinal arteriolar equivalent (CRAE) and central retinal venular equivalent (CRVE) were derived from the six largest arterioles and venules traversing Zone B. These metrics, calculated in pixels by the software, have been validated in previous studies[27,28]. To enhance interpretability, measurements were converted from pixels to microns using a conversion factor. This factor was determined based on an average adult optic disc diameter of 1800 µm[29] and the mean optic disc diameter across all included eyes.

In **Paper III**, we used DL and implemented a stepwise strategy including vessel segmentations, image feature extraction and additional numerical vessel features. This was done to provide the DL model with as much information as possible to predict PDR.

The first step in our approach involved extracting vessel segmentations from each input image using the FR-UNet model[30]. The FR-UNet is a DL model designed for medical and biological image segmentation tasks trained on vessel segmentation datasets, where the model learned to identify and extract vessel structures from raw images. We then added the vessel segmentation masks as

a fourth channel to the original retinal images. This enhanced the feature representation by incorporating vessel-specific information (Figure 7).

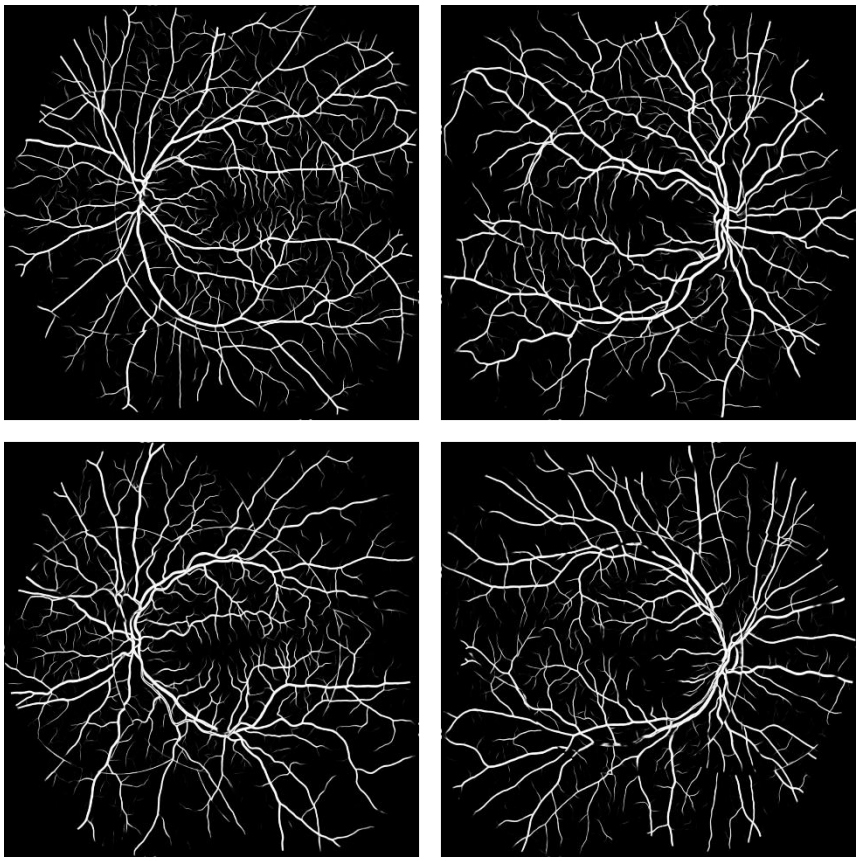


Figure 7: Four examples of skeletonized vessels from pre-processed retinal images with the FR-Unet model before training the deep learning prediction model. Adapted and modified from Paper III.

To extract deep features from the processed images (original images and vessel mask), we employed the InceptionV3 model that is a well-established CNN initially trained on ImageNet for large-scale image classification[31]. The InceptionV3 is able to capture both fine-grained and high-level spatial patterns.

Since the primary goal was to adapt InceptionV3 to our specific classification

task, we followed a transfer learning approach. Instead of training the entire network from the start, we froze the early layers (which capture general image features) and only fine-tuned the last layers with our images. This ensured that the model retained its powerful feature extraction capabilities while adapting to the vessel classification task.

We incorporated numerical features to the DL model extracted from the vessel segmentation mask in addition to image-based features, which included vessel density and tortuosity. The vessel density was calculated as the ratio of vessel pixels to total image pixels:

$$\text{Vessel density} = \frac{\text{Number of vessel pixels}}{\text{Total image pixels}}$$

To assess the vessel tortuosity, we first applied skeletonized mask to obtain a one-pixel-wide representation of the segmented vessels (Figure 7) and then analysed the skeleton to detect bifurcation points where vessels split into branches. We defined the vessel segments by tracing from each bifurcation point until the next bifurcation point, the vessel reached its endpoint or if it extended beyond a predefined maximum distance. We then computed the tortuosity for each segment as the ratio of segment length to Euclidean distance (straight-line between start and endpoint):

$$\text{Tortuosity} = \frac{\text{Segment length}}{\text{Euclidean Distance}}$$

Since we computed tortuosity for multiple vessel segments in each image, we needed to summarize these values before passing them to the model. We used maximum tortuosity, minimum tortuosity and average tortuosity of each image as numerical features. These numerical features were processed separately using a fully connected Multi-Layer Perceptron (MLP) feedforward neural network. Finally, the extracted features from InceptionV3 (image-based features) and the MLP (numerical features) were concatenated into a joint representation, enabling the model to integrate both visual and numerical information for the final classification.

The combined feature representation was passed through a final fully connected classification layer, which predicted the target class. The model was trained using Binary cross-entropy loss and optimized using Adam optimizer. Since the dataset used for training was imbalanced with one class (PDR group) having significantly fewer image samples. To address class imbalance, we

adjusted the loss function by assigning higher weights to the minority class, ensuring the model remained attentive to its correct classification.

Statistics

In **Paper II**, categorical data were analysed using the Chi-square test and are presented as numbers with percentages. Continuous data were analysed using the Mann-Whitney U test was applied and are presented as medians with interquartile ranges.

A multivariable logistic regression model was constructed to test for associations between independent clinical variables and PDR. We adjusted the model for age, sex, HbA1c, blood pressure, duration of diabetes, LDL, and triglycerides. To account for the inclusion of two eyes from some patients, we constructed the model as a mixed-effects regression with cluster robust standard errors to address the non-independence of observations within clusters. We assessed intergrader agreement by calculation intraclass correlation coefficients using a two-way mixed-effects model.

9.03 Results

In Paper II, the population consisted of 39 cases contributing 52 eyes and 89 controls contributing 107 eyes. The baseline differences between the two groups can be studied in Table A2 (appendix).

The results of the multivariable logistic regression analysis are shown in Figure 8, where we examined whether independent parameters predict progression to PDR. HbA1c (odds ratio (OR) 1.54 per 10 mmoL/moL; 95% CI: 1.15 to 2.07; p=0.004), triglyceride (OR 1.39 per 1 mmoL/L; 95% CI: 1.03 to 1.86; p=0.03) and the duration of diabetes (OR 1.09 per 1 year; 95% CI: 1.03 to 1.16; p=0.003) were statistic significant predictors of progression to PDR. Retinal vascular diameters (CRAE and CRVE) and all other clinical and demographic parameters were not predictive for PDR.

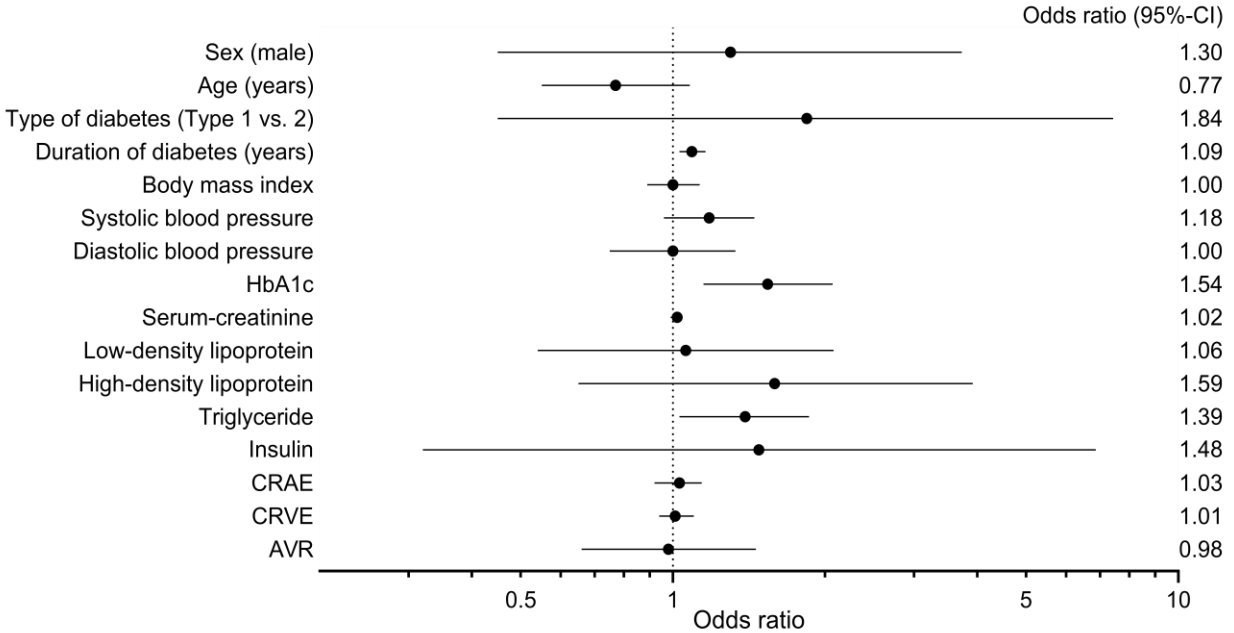


Figure 8: Multivariable logistic regression analysis with odds ratio for association between independent variables and proliferative diabetic retinopathy. CRAE and CRVE = central retinal arteriolar and venular equivalent; AVR = arterio-to-venule ratio. Adapted and modified from Table 2 in paper II.

For **Paper III**, we included 627 retinal images in the PDR case group and 26,201 images in the control group. The DL models with prediction from each individual DR level included 3,672 (ICDR level 1), 5,867 (ICDR level 2) and 1,693 (ICDR level 3) retinal images (Table 4).

Table 4: The Four Datasets and The Number of Images Dichotomized Into Training, Validation and Testing Datasets.

	All		**ICDR level 1**		**ICDR level 2**		**ICDR level 3**	
Subset, n	Control	Cases	Control	Cases	Control	Cases	Control	Cases
Training	19,812	462	2,746	17	4,373	146	1,001	284
Validation	2,594	95	352	2	556	39	116	54
Testing	3,795	70	552	3	730	23	200	38
Total, n	26,201	627	3,650	22	5,659	208	1,317	376

ICDR = International Classification of Diabetic Retinopathy Severity Scale. Adapted and modified from Paper III.

Despite the efforts in feature engineering, model architecture design, and handling class imbalance, the results indicate that the model is incapable of predicting the development of PDR resulting in zero true positive predictions as shown in Table 7.

Table 7: Performance for all of the Four Deep Learning Models for Predicting Future PDR.

	%
Sensitivity	0
Specificity	0
Positive predictive value	0
Negative predictive value	0

PDR = proliferative diabetic retinopathy. Adapted and modified from Paper III.

10 Paper IV

10.01 Objective

Our aim was to develop and validate a DL system able to identify patients with active PDR and distinguish these, not only from those with DR level 0 to 3, but also from those with stable inactive PDR (Figure 9). We aimed to achieve this through the annotation of new vessels and preretinal haemorrhages with DL. This approach will enable us to identify incident PDR, but also those already diagnosed with PDR in need of treatment and then distinguish these from those with inactive PDR stabilized by panretinal photocoagulation.

10.02 Methods

Study Design

This DL diagnostic accuracy study was conducted using retinal images from the Danish screening program for DR. The study took place in the Island of Funen, Denmark.

Study Population and Data Sources

The study population was defined within the FDDB from 2009 to 2019 and included retinal images from patients with type 1 or 2 diabetes. The CPR numbers, screening dates and ICDR level were extracted from FDDB while retinal images were retrieved from a local image database.

Inclusion and Exclusion Criteria

We included images representing DR level 4 from patients with a PDR diagnosis. Each patient could contribute multiple images but from different screenings episodes. The included retinal were of 6,528 x 6,528 pixels and were six-field mosaics. We included images of varying quality and excluded images if the retina was unrecognisable (five images excluded for being ungradable).

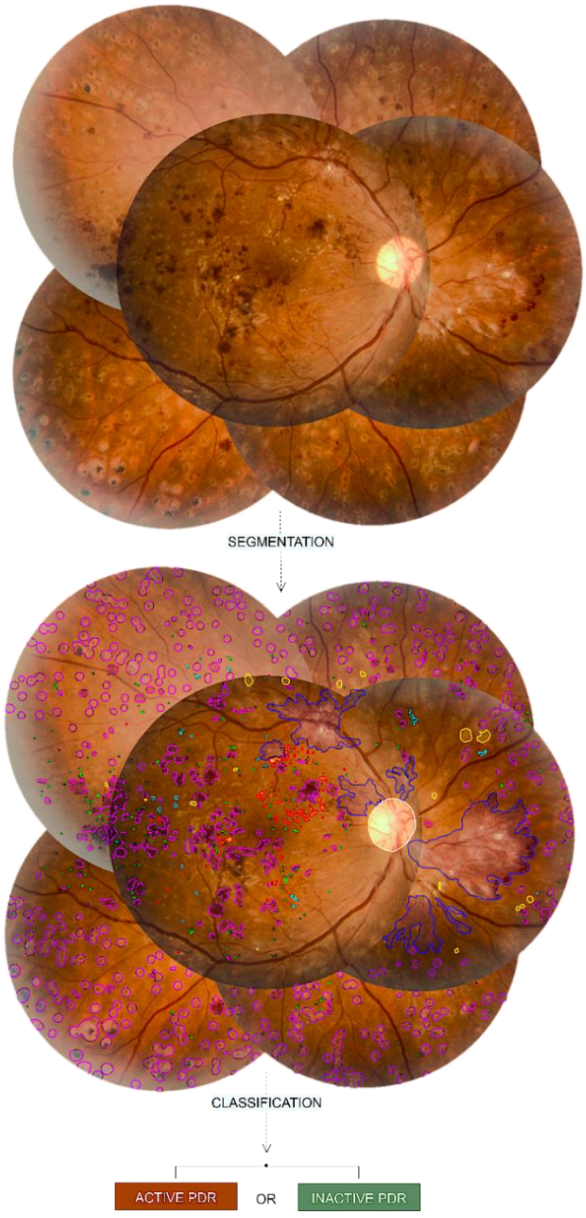


Figure 9: The figure illustrates the objective of Paper IV presenting an example of an original six-field retinal image and the same image with pre-segmentation of lesions from all ICDR levels 1 to 4. PDR = proliferative diabetic retinopathy. Adapted and modified from Paper IV.

Classification and Annotation of Retinal Images

The preparation of the retinal images began with a classification phase, followed by a DR lesion annotation phase to prepare the image data for training of the DL model. The grading team consisted of certified Grader 1 (SD) and two independent certified Graders 2 and 3 both with ophthalmological experience. All graders completed a DR grading course including video lessons and practical tasks to receive the VIOLA certification[32].

The initial process included categorizing the images as active or inactive PDR cases performed by Grader 1. Graders 2 and 3 reassessed the images to ensure correct classification. A joint assessment and decision was carried out in cases of disagreement between Grader 1 and one of the two remaining graders. This second classification changed 27 images from inactive to active PDR and 14 images from active to inactive PDR.

The annotation phase started with uploading the images to the Computer Vision Annotation Tool (CVAT) followed by applying our already developed DL segmentation model with the purpose of annotating nine different DR lesions[33]. These included microaneurysms, soft and hard exudates, intraretinal haemorrhages, intraretinal microvascular abnormalities (IRMA), peripheral and central laser tracks, preretinal haemorrhages and new vessels. Grader 1 conducted a manual review and refinement of the pre-segmented new vessels and preretinal haemorrhages to ensure optimal precision. This process also involved the segmentation of any missed preretinal haemorrhages or new vessels in the active PDR images as well as removing incorrect annotations from the pre-segmentation process. Additionally, Graders 2 and 3 evaluated the inactive PDR images to identify and delete incorrect annotations of preretinal haemorrhages and new vessels.

Algorithm Development

To develop the DL model to detect new vessels and preretinal haemorrhages, we allocated 70% of the retinal images for training, 10% for tuning and 20% for testing. It was made sure that a unique individual contributed images to only one of the allocations. The testing dataset also included retinal images of inactive PDR to evaluate model performance. All images had detailed annotations for new vessels, preretinal haemorrhages, and other relevant retinal features such as microaneurysms, retinal haemorrhages, hard exudates, cotton wool spots, IRMAs, peripheral plus central laser scars, and the optic disc (Figure A2). All of these were generated by our previous developed segmentation model[33]. We simplified the labels into four categories for training purposes that included background, new vessels, preretinal haemorrhages, and other lesions.

Our segmentation model is based on a UperNet[34] architecture with a Con-

vNext[31] backbone that was pre-trained on ImageNet. The training involved 440 images of active PDR with 1,381 manually annotated new vessels and preretinal haemorrhages, while the tuning included the use of 63 images with 123 annotations. The model was implemented in PyTorch in which we utilized the TorchSeg library[35]. The output was established by using convolutional layers of different kernel sizes (1×1, 3×3, and 5×5) to enhance the ability to detect features at various levels of detail. We merged the final probability maps from these layers into a single probability map.

We applied a range of techniques to improve the models ability to recognize and detect small or rare lesions as well as performing under different image conditions. These included to handle class imbalance with the application of a loss function as a combination of Log Cosh Dice Loss[36] and Focal Tversky Loss[37]. It also included modification of the images to add variety with augmentation, which rotated and brightness adjusted the images, to help the model recognize lesions under different conditions[38]. The six-field retinal images were processed in overlapping 288×288 pixel patches with smaller 256×256 pixel crops randomly sampled during training. Additionally, we slightly modified the lesion labels with morphological transformations to account for small uncertainties in manual annotations to improve accuracy.

We prioritized the patches that contained new vessels and preretinal to prevent bias toward background regions during training. We gradually increased background patch inclusion as the training progressed. The Adam optimizer with a cyclical learning rate fine-tuned the model and we selected the best-performing weights based on the Dice coefficient score.

An initial test of the DL models ability to detect new vessels and preretinal haemorrhages revealed a sensitivity of 100% and a low positive predictive value (PPV) of 12%. The model obviously produced many false positive segmentations for new vessels and/or preretinal haemorrhages, which had us carry out a qualitative review of the output images. We identified to main errors, with the first being small dot-like detections in random areas (compared to the reference) and the second being confusion with irregular vascular structures like IRMAs. We introduced a threshold to ignore detected regions smaller than 10,576 pixels (0.02% of the total retinal image area) to reduce the main error type, which was small and random dot-like segmentations.

Statistics

We used STATA17 to calculate sensitivity, specificity, PPV and negative predictive value (NPV) and their respective 95%-confidence intervals.

10.03 Results

We identified 3,263 PDR images divided as 638 with active PDR and 2,624 with inactive PDR. All of the 638 active PDR images were included and allocated for training (440 images), tuning (63 images) and testing (135 images) as shown in Table 6. The training dataset contained of 1,381 new vessel and preretinal haemorrhage lesions, the tuning dataset of 123 and the testing dataset of 374. The testing dataset also included 301 images with inactive PDR.

Table 6: Numbers of Images and Segmentation Counts for Preretinal Haemorrhages and New Vessels across Dataset Splits

	Patients, n	Retinal images, n	Active PDR images, n	New vessels and preretinal haemorrhages, n
Training (70%)	105	440	440 (68.9)	1,381 (73.5)
Tuning (10%)	29	63	63 (9.9)	123 (6.5)
Testing (20%)	65	436	135 (21.2)	374 (19.9)
Total	199	939	638	1,878

PDR = proliferative diabetic retinopathy. Adapted and modified from Paper IV.

The confusion matrix (Figure 10) represents the 426 images included in the testing set, which revealed 121 true positives, 211 true negatives, 91 false positives and 13 false negatives.

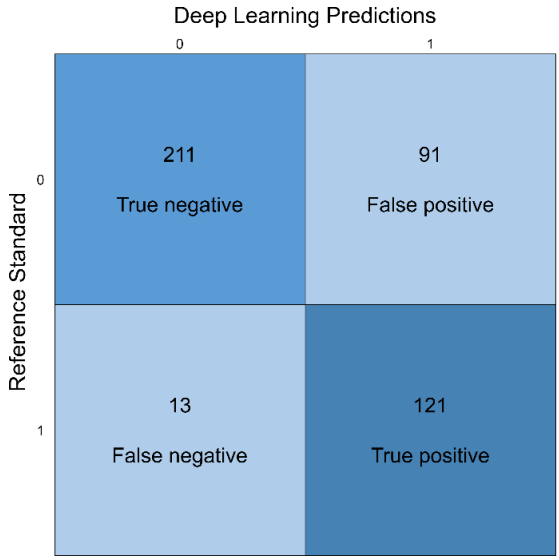


Figure 10: The confusion matrix shows the number of true negative, true positive, false negative and true negative predictions by the deep learning model compared to the reference standard. Adapted and modified from paper IV.

The DL model achieved a sensitivity of 90% (95%-CI: 85% to 95%), a specificity of 70% (95%-CI: 65% to 75%), a PPV of 57% (95%-CI: 50% to 64%) and a NPV of 94% (95%-CI: 91% to 97%).

Table 7: Diagnostic Performance of the Deep Learning Model for Detecting Active PDR at the Image Level

	% (95% CI)
Sensitivity	90 (85 to 95)
Specificity	70 (65 to 75)
Positive predictive value	57 (50 to 64)
Negative predictive value	94 (91 to 97)

The definition of active PDR is the presence of one or more preretinal haemorrhages and/or new vessels. The table shows diagnostic accuracy at the 0.9 lesion threshold of 10,576 pixel (0.02% of the total image area). The reference is the established ground truth by the three certified graders. PDR = proliferative diabetic retinopathy. 95% CI = 95% confidence interval.

Figure 11 shows an output example of our proposed model that consists of a full mask overview and a zoomed region in the same image. The zoomed region shows two larger new vessels segmented by the DL model with a high precision while it also exemplifies the dot-like error type as well as the segmentation of a questionable irregular vascular formation in the right site of the image.

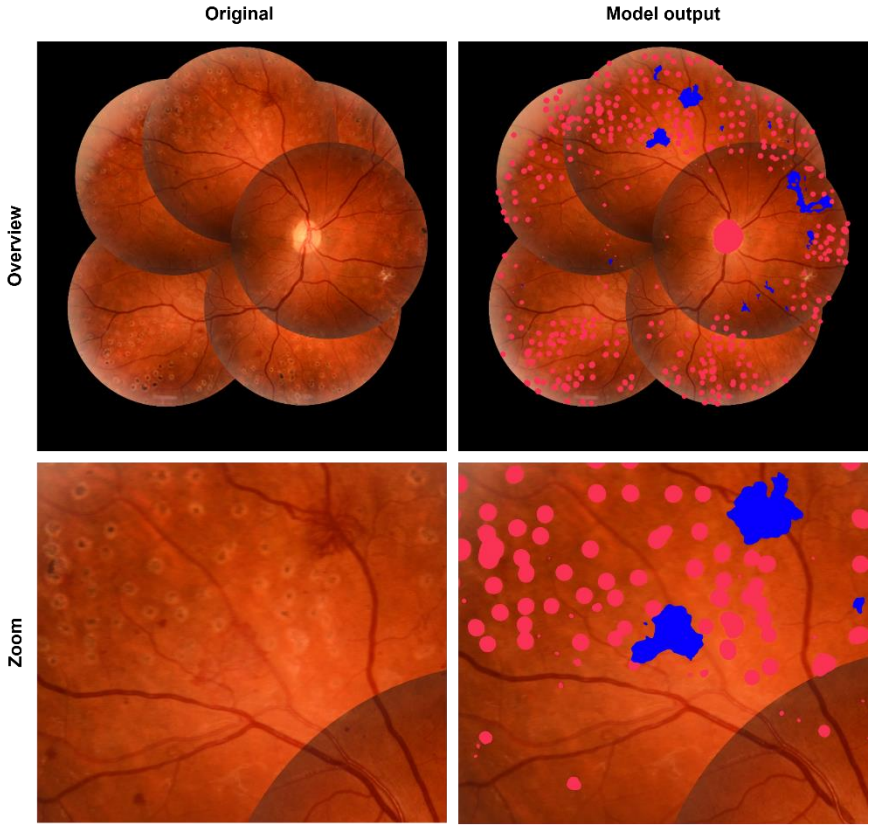


Figure 11: An output image showing the annotations of new vessels by the deep learning segmentation model. The blue annotations are new vessels and all other annotations is in pink colour. Adapted and modified from paper IV.

11 Discussion

This thesis investigated the detection and progression of PDR with the use of DL, epidemiological analysis, and vascular imaging. The epidemiological study of a nationwide diabetes cohort revealed much lower PDR incidence rates as compared to earlier studies. The thesis confirmed the association between key risk factors and progression to PDR, including systemic comorbidity, type 1 diabetes, use of insulin, triglycerides, HbA1c and longer duration of diabetes. Retinal vascular calibres did not predict PDR development from early retinal images with no or very minimal DR, but we achieved a sensitivity of 90% and a specificity of 70% with our DL model to detect active PDR.

11.01 Methodological considerations

Retinal Vessel Geometry Analysis

Changes in retinal vessel diameters are described as the first detectable changes in DR. CRAE and CRVE have been extensively investigated in relation to vascular health, with research indicating wider venules and narrower arterioles are associated with progression of DR and the development of PDR[39]. However, these findings are based on eyes already in advanced disease stages (ICDR levels 2 and 3). Given our unique dataset of retinal images from individuals with no or mild DR and a long follow-up time, we aimed to investigate if CRAE or CRVE at this early stage could predict future PDR. If we could identify patients at risk before they progress to more severe stages of DR (ICDR 2 and 3) it would be highly valuable and would allow early optimization of all modifiable factors to prevent further disease progression.

This study was essential because no previous research has attempted to predict PDR from CRAE and CRVE at such an early stage of DR. Investigating this measure in isolation provides valuable insight into disease pathology and helps clarify its role in progression. Moreover, if CRAE and CRVE proved to be a strong predictive factor, it could be integrated into a DL prediction model for PDR alongside other known risk parameters, such as HbA1c, to improve long-term prediction and early disease prevention.

Deep Learning for Existing Disease Detection

Manual data preparation to establish a reference standard before algorithm training is the most critical phase in diagnostic studies[40]. To enhance label reliability, we incorporated an adjudication process in image preparation, as it has shown to reduce errors in DR classification[40]. This process included an initial classification (Grader 1), independent secondary classifications (Grader 2 and 3), identification of disagreements and a final consensus review to determine the reference label. While this approach is robust, there is always room for optimization. First, retinal specialists are preferred for classification tasks, but the availability as a recourse is sparse. Secondly, the challenge to classify active versus inactive PDR often arises from distinguishing between IRMA versus new vessels. The distinction can be challenging in the clinic, where a fluorescein angiography is conducted in cases of doubt. This IRMA versus new vessel issue also contributed to false positive predictions by the DL model. Label optimization could be improved through manual review of patients' medical records for clarification.

It is important to recognize the DL model as a segmentation model and not as a classification model. We chose to train the model from active PDR images alone, which can seem illogical since we aimed to distinguish active from inactive PDR. The training set should be regarded as a composite of various lesions from all included images, encompassing new vessels and preretinal haemorrhages that define active PDR, as well as lesions characteristic of other DR levels, such as microaneurysms, soft exudates, intraretinal haemorrhages and so on. We therefore considered this approach efficient to fulfil the purpose to distinguish active from inactive PDR. Even though we only used active PDR images, new vessels and preretinal haemorrhages are still underrepresented compared to all other lesions, which is why we upsample (and undersample inactive images) these to a certain degree before overfitting the model[41]. However, the specificity could be better and it should be considered if retraining of the model with a number of inactive PDR image could improve this performance metric.

Careful consideration is required when modifying the DL model post hoc to exclude lesions smaller than 10,576 pixels. The purpose was to balance the model for clinical implementation by improving the specificity and prioritize fewer false positive predictions at the cost of some missed active PDR predictions (false negatives). We set the threshold to a sensitivity of 90% as this corresponds an acceptable level of diagnostic accuracy compared to human graders reported to range from 82.7% to 88.5% in other studies[42-44].

It is important to determine whether the ignored false-positive lesions are unlikely to have a clinical impact. The excluded lesions underwent qualitative re-

view by human graders, but to ensure patient safety, it is essential to establish evidence that early-stage disease is not being missed through follow-up evaluation of the relevant cases. However, the NPV of 94% provides confidence in the model's ability to identify cases without disease.

Deep Learning for Long-Term Outcome Prediction

Several studies have successfully predicted various outcomes from retinal images. Poplin et al.[45] reported an area under the curve (AUC) of 0.97 for sex prediction and 0.71 for smoking status while Sabanayagam et al reported an AUC of 0.84 for chronic kidney disease prediction[46]. These all highlight the potential of AI as a diagnostic tool, but common for these studies was that they predicted existing sex, smoking status and disease status and not future outcomes. Zhang et al., on the other hand, reported an AUC of 0.83 for predicting chronic kidney disease four years before onset[47]. This was achieved with a DL model trained solely on retinal images, which included a total of 60,244 images of which 10,977 (18.2%) were cases. A study by Dai et al[48] predicted both time to DR progression and onset of vision-threatening DR (ICDR level 3 to 4) from non-vision-threatening DR in one to five years with AUCs reaching 0.77 to 0.86. The model was developed from nearly a million fundus images used for pretraining, development and internal validation of their DL model. This emphasizes the need for large datasets as we did not near us the amount of case images (or total images) compared to these studies[47,48]. We tried to feed our model as much information as possible by adding vessel segmentations, automatic image feature extraction and additional numerical vessel features. We further tried to predict from increasing DR levels without success. Future model development should include demographic and clinical data in combination with the retinal images to provide the DL model as much relevant information as possible.

11.02 Clinical Implications

The current Danish national DR screening system is effective, but many patients screened show no or minimal signs of DR pathology with the result of unnecessary strain on the screening system and highlights the need for efficient resource allocation. Epidemiological analysis is essential for understanding disease patterns in large populations, to identify risk factors and to ensure that any alterations are evidence-based. Of the technological approaches investigated in this thesis, DL is by far the most promising for implementation. Retinal geometric analysis using semi-automatic software like VAMPIRE is, first, too time-consuming to be a practical solution in a clinical setting, second, results

vary across studies and different platforms[49] and, third, if retinal vessel geometry is important for DL predictions, the model will inherently learn these features, which will make separate geometric analysis unnecessary.

The implementation of a DL system for classification of vision-threatening DR (active PDR) is not intended a standalone solution for the Danish DR screening program, where there are sufficient medical resources to take care of patients at earlier DR stages. However, a targeted DL system to accurately and early detect active PDR is still essential to prioritize patients requiring urgent treatment to reduce risk of sight loss. Beyond its potential for clinical implementation in Denmark, the proposed model could address the challenge of limited healthcare resources in low- and middle-income countries, where insufficient DR screening highlights a significant healthcare inequity[1]. DL systems with lower threshold for referral to an ophthalmologist would not be beneficial in low- to middle income countries as the medical resources needs to be allocated to those in need of treatment.

Screening Intervals

We showed by investigating 201,945 patients with diabetes (*Paper I*) that only 0.1% with ICDR level 0 had progression to PDR in a five-year period. This finding, among others[50], have resulted in prolonged screening intervals to a minimum of 24 months for patients with no DR in the Danish screening program. This is probably mostly due to advancements in diabetes management as described in the background section, however, the screening program itself also plays a crucial role as it enables early detection and timely intervention. Together, these has reduced the proportion of blindness in the Danish diabetes population by 78%[51].

Deep Learning: From Bench to Bedside

Many DL models have demonstrated strong performance, but their clinical implementation remains limited. There are several steps a diagnostic DL model must go through before implementation, including data collection and preprocessing, model development, internal and external testing, regulatory approval from healthcare authorities, clinical validation from prospective trials and finally implementation, which require staff training, monitoring and continuous improvements.

Several challenges arise at different stages with dataset homogeneity being an issue that impairs a model's generalization potential[52]. It is important to diversify the dataset in terms of ethnicity, image quality and image modality to address this issue[52].

Another limitation is the reliance on small datasets during development, which is particularly problematic for rare diseases. For example, the commercial algorithm developed by Gulshan et al.[53] was trained on a dataset where only 3.7% of images contained PDR and did not distinguish between active and inactive disease. While humans can learn to recognize features from only a few examples, DL algorithms require large datasets to achieve the same level of proficiency[54]. Although we provide a larger number of PDR images in *Paper IV* compared to previous studies[42,55], our algorithm still faces challenges associated with limited data.

Several methods have been presented to deal with the small data problem. A straightforward method is to balance the data, which involves to upsample the rare classes (active PDR cases) and downsample the common classes (inactive PDR cases)[56]. However, upsampling, as used in our DL model in *Paper IV*, increases the risk of overfitting, making the model less generalizable to new datasets. We used data augmentation to generate additional training data to minimize the risk of this issue, which was done by random alterations of the training data[54]. Another approach was the use of different kernel sizes to enhance the models ability to extract important features. This technique preprocesses the images at multiple scales to capture fine details as well as broader structural patterns. By applying varying kernel sizes, smaller kernels focus on local textures and edges, while larger kernels capture more global contextual information[57].

The black box phenomenon is a common concern among clinicians as the lack of interpretability in traditional black box models can undermine trust and hinder integration into clinical practice. It lies in the nature of health care professionals that they seek to understand how decisions are made to ensure patient safety. Therefore, several approaches to open the black box has been proposed. Heatmapping is one of the most commonly used in image recognition, which allows visualization of the areas in an image that the DL model focuses on to make its predictions[22]. However, the highlighted areas are often difficult for healthcare professionals to interpret, as the model may predict from features beyond human understanding. We chose to develop a DL segmentation model (*Paper IV*) constrained to features within the realm of human comprehension to address this problem. By aligning our DL model with human understanding of DR, we developed a segmentation model that provides accurate visual annotations of new vessels and preretinal haemorrhages, enhancing clinical decision-making and eliminating the black box. This eases the clinical implementation, but there are still some concerns to be addressed. Yu et al [58] reported that errors from an AI assistance model for X-ray diagnostic tasks strongly affected clinicians' decision-making and influenced treatment outcomes. It underscores the need for

high-quality DL models and prospective clinical monitoring of treatment outcomes when integrating AI into DR screening.

11.03 Limitations

Paper I

The unavailability of important information such as glycaemic control, blood pressure, BMI, smoking status, and physical activity limits the ability to fully adjust for important risk factors. Second, data on diabetic macular oedema were inconsistently reported, and we therefore had to use anti-VEGF injections as a proxy measure, which may not accurately reflect its presence or severity. Third, [59]while the study utilized high-quality national registries, variability in data entry and intergrader differences in DR classification could introduce misclassification bias. Additionally, the observational design prevents causal inferences, and residual confounding cannot be excluded despite statistical adjustments. Finally, the study population consisted of individuals who attended the national DR screening, and we can therefore not rule out if selection bias is introduced by excluding those who do not participate in the DR screening program.

Paper II and III

The major limitation of these studies was the lack of sufficient data to confidently conclude that PDR cannot be predicted from retinal images alone. It was a limitation that Paper II only evaluated vascular diameters as other measures such as vessel tortuosity, density and fractal dimension would have been of high value to investigate. We chose to exclude these measures as the overlapping edges of the six-field images made the analysis with VAMPIRE difficult to replicate for other researchers. Vessel diameter was the only measure not conflicting with the edges of the six-field images. Similarly, we found no suitable method for quantification of vascular diameters to be implemented in the DL model. The class imbalance caused by the limited number of PDR cases and in Paper III weakens the foundation for DL model development, as the model lacks sufficient exposure to features characteristic of future PDR progression needed to predict a rare outcome like PDR. Additionally, the total number of retinal images included is a limitation, as studies reporting successful results have used substantially larger datasets to improve feature extraction[48,59].

Paper IV

A major limitation is the lack of external validation, which leaves us uncertain about its ability to generalize to other populations. Furthermore, class imbal-

ance, along with the challenges posed by data limitations, increases the risk of overfitting. This is a recurring challenge when investigating rare diseases, yet despite this, we achieved promising results as a foundation for future refinements. The population mainly consisted of Caucasian individuals, which emphasizes the importance of validating the algorithm on external multi-ethnic cohorts to ensure the model can be safely implemented without ethnic bias.

11.04 Conclusions

This thesis presents a comprehensive investigation into PDR with the use of DL segmentation, epidemiological assessment, and vascular imaging. While DL effectively detected existing disease, neither DL nor retinal vascular imaging predicted incident PDR. We identified key risk factors associated with PDR progression and confirmed a decreasing trend in PDR incidence rates, particularly for well-regulated populations.

The developed DL segmentation model demonstrated a high sensitivity (90%) and an acceptable specificity (70%) to distinguish active PDR from inactive PDR cases. The DL segmentations model's ability to annotate new vessels and preretinal haemorrhages strongly enhances transparency and clinical decision-making. Guidance from segmentations prior to training also appears to enhance a DL model's ability to extract important features. However, this approach is not feasible when clinicians do not know which features to target in the retinal images, as is the case when predicting future PDR.

The findings from these studies underscores the potential for integrating DL models into DR screening programs while maintaining a strong focus on systemic disease management.

12 Perspectives

These findings contribute to our understanding of the current status of PDR in a Danish nation-wide population. We were encouraged to discover the low incidence rates of PDR in this well-regulated low-risk population, but the need for continuous DR screening should to be highlighted, as the prevalence of diabetes is projected to increase worldwide. Future large scale cohort studies in the next few decades is of huge importance to continuously evaluate the PDR status.

The workload on the screening program for DR is expected to increase as the prevalence of diabetes increases and changes in the demography with a growing proportion of older individuals. This underscores the importance of working on innovative methods to reduce the workload on the healthcare system. DL presents great potential in reducing the workload, but it is important that any given model meets the quality standards we must uphold to ensure patient safety.

The predictive potential of DL must be further explored despite our failure to predict PDR in this thesis. Incorporating additional risk factors into a DL model could enhance its performance given their crucial role in PDR progression.

13 Dansk Resumé

På verdensplan lever over 828 millioner mennesker med diabetes. Diabetisk nethindesygdom er en velkendt komplikation til diabetes og udvikles hos mange patienter i forskellig grad efter længere tid med forhøjet blodsukker. Diabetisk nethindesygdom forløber over fem stadier startende med ingen nethindeforandringer (niveau 0), minimale nethindeforandringer (niveau 1), moderate nethindeforandringer (niveau 2), svære ikke synstruende nethindeforandringer (niveau 3) og et synstruende slutstadium kaldet proliferativ diabetisk nethindesygdom (niveau 4).

Diabetisk nethindesygdom skyldes sygdom i nethindens små blodkar, hvori der opstår dysfunktion, der fører til iltmangel i nethinden og som en konsekvens heraf dannelse af nye skrøbelige blodkar. De nye skrøbelige blodkar har tendens til at bløde ud i øjets glaslegeme og udgør en reel risiko for synstab. Det er vigtigt at diagnosticere disse, så de kan behandles før et blivende synstab opstår. Patienter med diabetes indkaldes derfor til øjenscreening med passende intervaller bedømt ud fra deres niveau af nethinde sygdom og langtidsblodsukker måling.

Ph.d. afhandlingen havde til formål at undersøge diabetisk nethindesygdom fra flere perspektiver, herunder en epidemiologisk tilgang samt udvikling af diagnostiske metoder til at identificere og forudsige udviklingen af det synstruende slutstadie proliferativ diabetisk nethindesygdom.

I Studie I udførte vi en epidemiologisk analyse af en national population på 201.945 diabetes patienter, hvor vi undersøgte nuværende antal af patienter med proliferativ diabetisk nethindesygdom og hvor mange der udviklede det synstruende slutstadium i løbet af et, tre og fem år. Studie II og III til formål at forudsige fremtidig udvikling af proliferativ diabetisk nethindesygdom analyseret ud fra fotos af nethinden taget før diagnosetidspunktet. Vi undersøgte om diameteren for nethindens kar, analyseret med et semiautomatisk softwareprogram, kunne benyttes til at forudsige udviklingen af proliferativ nethinde sygdom og om en kunstig intelligens model udviklet fra nethinde fotos kunne benyttes til samme formål. I Studie IV udviklede vi en kunstig intelligens model til diagnostik af synstruende nethindesygdom ud fra nethinde fotos, hvor modellen skulle lære at identificere og indtegne skrøbelige nye blodkar og blødninger ud i øjets glaslegeme fra de nydannede blodkar.

I den første undersøgelse fandt vi centrale risikofaktorer for udvikling af proliferativ diabetisk nethinde sygdom, hvor færre udvikler tilstanden sammenlignet

med tidligere tiders opgørelser. I studie II og III fandt vi at proliferativ diabetisk nethindesygdom ikke kunne forudsiges ud fra de metoder vi benyttede og de data vi havde til rådighed. Kunstig intelligens modellen i det fjerde og sidste studie kunne med stor sikkerhed identificere synstruende nethindesygdom ved at diagnosticere og tilmed indtegne nye skrøbelige blodkar og blødninger herfra.

Fundende for indeværende ph.d. afhandling bidrager med et værdifuldt indblik i den nuværende status for den proliferative diabetiske nethindesygdom. På trods af færre nye tilfælde af patienter, der når det synstruende slutstadium er lovende, understreger den globalt voksende gruppe af patienter der diagnosticeres med diabetes, at behovet for screening fortsat er tilstede. Kunstig intelligens er en lovende metode til at håndtere den forventede øgning i arbejdsbyrden i takt med at flere patienter oplever at få diabetes. Der er behov for mere forskning, for at indfri potentialet til at forudsige proliferativ diabetisk nethinde sygdom med kunstig intelligens, hvor de primære risikofaktorer fra Studie I med fordel kan inddrages, for at forbedre modellens nøjagtighed.

14 References

1. Vujosevic S, Aldington SJ, Silva P, et al. Screening for diabetic retinopathy: new perspectives and challenges. *Lancet Diabetes Endocrinol*. Apr 2020;8(4):337-347. doi:10.1016/s2213-8587(19)30411-5
2. Saeedi P, Petersohn I, Salpea P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. *Diabetes Res Clin Pract*. Nov 2019;157:107843. doi:10.1016/j.diabres.2019.107843
3. Zhou B, Rayner AW, Gregg EW, et al. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. *The Lancet*. 2024;404(10467):2077-2093. doi:10.1016/S0140-6736(24)02317-1
4. Antonetti DA, Klein R, Gardner TW. Diabetic retinopathy. *N Engl J Med*. Mar 29 2012;366(13):1227-39. doi:10.1056/NEJMra1005073
5. Teo ZL, Tham YC, Yu M, et al. Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045: Systematic Review and Meta-analysis. *Ophthalmology*. Nov 2021;128(11):1580-1591. doi:10.1016/j.ophtha.2021.04.027
6. Yau JW, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*. Mar 2012;35(3):556-64. doi:10.2337/dc11-1909
7. Abou Taha A, Dinesen S, Vergmann AS, Grauslund J. Present and future screening programs for diabetic retinopathy: a narrative review. *Int J Retina Vitreous*. Feb 3 2024;10(1):14. doi:10.1186/s40942-024-00534-8
8. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. Glycosylated hemoglobin predicts the incidence and progression of diabetic retinopathy. *Jama*. Nov 18 1988;260(19):2864-71.
9. Gabbay KH. Editorial: Glycosylated hemoglobin and diabetic control. *N Engl J Med*. Aug 19 1976;295(8):443-4. doi:10.1056/nejm197608192950810
10. Nordwall M, Bojestig M, Arnqvist HJ, Ludvigsson J. Declining incidence of severe retinopathy and persisting decrease of nephropathy in an unselected population of Type 1 diabetes—the Linköping Diabetes Complications Study. *Diabetologia*. 2004/07/01 2004;47(7):1266-1272. doi:10.1007/s00125-004-1431-6
11. O'Leary F, Campbell M. The blood-retina barrier in health and disease. *Febs j*. Feb 2023;290(4):878-891. doi:10.1111/febs.16330

12. Photocoagulation treatment of proliferative diabetic retinopathy: the second report of diabetic retinopathy study findings. *Ophthalmology*. Jan 1978;85(1):82-106. doi:10.1016/s0161-6420(78)35693-1
13. Principles and practice of screening for disease. *J R Coll Gen Pract*. 1968:318. vol. 4.
14. Grauslund J, Andersen N, Andresen J, et al. Evidence-based Danish guidelines for screening of diabetic retinopathy. *Acta Ophthalmol*. Dec 2018;96(8):763-769. doi:10.1111/aos.13936
15. National retningslinje for screening af diabetisk retinopati Guidelines. *Dansk Oftalmologisk Selskab*. 13. august 2024 doi:https://dansk-oftalmologisk-selskab.dk/nuvaerende-guidelines/
16. Wilkinson CP, Ferris FL, 3rd, Klein RE, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. Sep 2003;110(9):1677-82. doi:10.1016/s0161-6420(03)00475-5
17. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. May 28 2015;521(7553):436-44. doi:10.1038/nature14539
18. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60(6):84–90. doi:10.1145/3065386
19. Hinton G, Deng L, Yu D, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*. 2012;29(6):82-97. doi:10.1109/MSP.2012.2205597
20. Sutskever I. Sequence to Sequence Learning with Neural Networks. *arXiv preprint arXiv:14093215*. 2014;
21. Lillicrap TP, Santoro A, Marris L, Akerman CJ, Hinton G. Backpropagation and the brain. *Nat Rev Neurosci*. Jun 2020;21(6):335-346. doi:10.1038/s41583-020-0277-3
22. Petch J, Di S, Nelson W. Opening the Black Box: The Promise and Limitations of Explainable Machine Learning in Cardiology. *Can J Cardiol*. Feb 2022;38(2):204-213. doi:10.1016/j.cjca.2021.09.004
23. Andersen N, Hjortdal J, Schielke KC, et al. The Danish Registry of Diabetic Retinopathy. *Clin Epidemiol*. 2016;8:613-619. doi:10.2147/clep.S99507
24. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health*. Jul 2011;39(7 Suppl):30-3. doi:10.1177/1403494811401482
25. Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. *Scand J Public Health*. Jul 2011;39(7 Suppl):38-41. doi:10.1177/1403494810394717
26. Larsen MB, Henriksen JE, Grauslund J, Peto T. Prevalence and risk factors for diabetic retinopathy in 17 152 patients from the island of

Funen, Denmark. *Acta Ophthalmol*. Dec 2017;95(8):778-786. doi:10.1111/aos.13449
27. Lupaşcu CA, Tegolo D, Trucco E. Accurate estimation of retinal vessel width using bagged decision trees and an extended multiresolution Hermite model. *Med Image Anal*. Dec 2013;17(8):1164-80. doi:10.1016/j.media.2013.07.006
28. McGrory S, Taylor AM, Pellegrini E, et al. Towards Standardization of Quantitative Retinal Vascular Parameters: Comparison of SIVA and VAMPIRE Measurements in the Lothian Birth Cohort 1936. *Transl Vis Sci Technol*. Mar 2018;7(2):12. doi:10.1167/tvst.7.2.12
29. Jonas JB, Gusek GC, Naumann GO. Optic disc, cup and neuroretinal rim size, configuration and correlations in normal eyes. *Invest Ophthalmol Vis Sci*. Jul 1988;29(7):1151-8.
30. Liu W, Yang H, Tian T, et al. Full-Resolution Network and Dual-Threshold Iteration for Retinal Vessel and Coronary Angiograph Segmentation. *IEEE J Biomed Health Inform*. Sep 2022;26(9):4623-4634. doi:10.1109/jbhi.2022.3188710
31. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016; https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.308
32. Andersen JKH, Hubel MS, Savarimuthu TR, Rasmussen ML, Sørensen SLB, Grauslund J. A digital online platform for education and certification of diabetic retinopathy health care professionals in the Region of Southern Denmark. *Acta Ophthalmol*. Aug 2022;100(5):589-595. doi:10.1111/aos.15123
33. Andersen JKH, Hubel MS, Rasmussen ML, Grauslund J, Savarimuthu TR. Automatic Detection of Abnormalities and Grading of Diabetic Retinopathy in 6-Field Retinal Images: Integration of Segmentation Into Classification. *Transl Vis Sci Technol*. Jun 1 2022;11(6):19. doi:10.1167/tvst.11.6.19
34. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Springer International Publishing; 2015:234-241.
35. PyTorch. https://github.com/isaaccorley/torchseg
36. Jadon S. A survey of loss functions for semantic segmentation. 2020:1-7.
37. Abraham N, Khan NM. A Novel Focal Tversky Loss Function With Improved Attention U-Net for Lesion Segmentation. 2019:683-687.
38. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: Fast and Flexible Image Augmentations. *Information*. 2020;11(2):125.

39. Wong TY. Retinal vessel diameter as a clinical predictor of diabetic retinopathy progression: time to take out the measuring tape. *Arch Ophthalmol*. Jan 2011;129(1):95-6. doi:10.1001/archophthalmol.2010.347
40. Krause J, Gulshan V, Rahimy E, et al. Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy. *Ophthalmology*. Aug 2018;125(8):1264-1272. doi:10.1016/j.ophtha.2018.01.034
41. Buda M, Maki A, Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw*. Oct 2018;106:249-259. doi:10.1016/j.neunet.2018.07.011
42. Ting DSW, Cheung CY, Lim G, et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *Jama*. Dec 12 2017;318(22):2211-2223. doi:10.1001/jama.2017.18152
43. Chia MA, Hersch F, Sayres R, et al. Validation of a deep learning system for the detection of diabetic retinopathy in Indigenous Australians. *Br J Ophthalmol*. Jan 29 2024;108(2):268-273. doi:10.1136/bjo-2022-322237
44. Dai L, Wu L, Li H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nat Commun*. May 28 2021;12(1):3242. doi:10.1038/s41467-021-23458-5
45. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. Mar 2018;2(3):158-164. doi:10.1038/s41551-018-0195-0
46. Sabanayagam C, Xu D, Ting DSW, et al. A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. *Lancet Digit Health*. Jun 2020;2(6):e295-e302. doi:10.1016/s2589-7500(20)30063-7
47. Zhang K, Liu X, Xu J, et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat Biomed Eng*. Jun 2021;5(6):533-545. doi:10.1038/s41551-021-00745-6
48. Dai L, Sheng B, Chen T, et al. A deep learning system for predicting time to progression of diabetic retinopathy. *Nat Med*. Feb 2024;30(2):584-594. doi:10.1038/s41591-023-02702-z
49. Dinesen S, Stokholm L, Subhi Y, et al. Retinal main vessel calibers and systemic markers for long-term development of proliferative diabetic retinopathy. *Acta Ophthalmol*. Jun 2024;102(4):448-454. doi:10.1111/aos.15780
50. Grauslund J, Stokholm L, Andersen N, et al. Risk of 2-year progression from no diabetic retinopathy to proliferative diabetic retinopathy in accordance with glycaemic regulation in 145 527 persons with diabetes in a

national cohort. *Acta Ophthalmologica*. 2024;102(3):e407-e409. doi:https://doi.org/10.1111/aos.15806
51. Blindbaek S SL, Heegaard S et al. 31st Meeting of the European Association for the Study of Diabetes Eye Complications Study Group (EASDec) Odense, Denmark, 28th – 30th October 2021. *European Journal of Ophthalmology*. 2021;31(2_suppl):1-41. doi:10.1177/11206721211047031
52. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol*. Feb 2019;103(2):167-175. doi:10.1136/bjophthalmol-2018-313173
53. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *Jama*. Dec 13 2016;316(22):2402-2410. doi:10.1001/jama.2016.17216
54. Rather IH, Kumar S, Gandomi AH. Breaking the data barrier: a review of deep learning techniques for democratizing AI with small datasets. *Artificial Intelligence Review*. 2024/08/02 2024;57(9):226. doi:10.1007/s10462-024-10859-3
55. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med*. 2018;1:39. doi:10.1038/s41746-018-0040-6
56. Shen L, Lin Z, Huang Q. Relay Backpropagation for Effective Learning of Deep Convolutional Neural Networks. 2015:
57. Kang G, Dong X, Zheng L, Yang Y. PatchShuffle Regularization. *ArXiv*. 2017;abs/1707.07103
58. Yu F, Moehring A, Banerjee O, Salz T, Agarwal N, Rajpurkar P. Heterogeneity and predictors of the effects of AI assistance on radiologists. *Nature Medicine*. 2024/03/01 2024;30(3):837-849. doi:10.1038/s41591-024-02850-w
59. Bora A, Balasubramanian S, Babenko B, et al. Predicting the risk of developing diabetic retinopathy using deep learning. *The Lancet Digital Health*. 2021;3(1):e10-e19. doi:10.1016/S2589-7500(20)30250-8

15 Appendix

Table A1 (Paper I): Baseline characteristics of patients in the Danish Registry of Diabetic Retinopathy according to the level of diabetic retinopathy in the worst eye.

	Overall	**Level 0**	**Level 1**	**Level 2**	**Level 3**	**Level 4**
Number of patients, n	201,945	171,790	21,120	6,583	1,159	1,293
Sex, n (%) male	114,107 (56.5)	95,842 (55.8)	12,517 (59.3)	4,156 (63.1)	786 (67.8)	806 (62.3)
Age, years (IQR)	65.8 (55.6;73.2)	66.2 (56.3;73.4)	63.5 (51.5;72.2)	62.0 (51.4;70.7)	56.0 (45.6;66.0)	64.1 (52.2;72.0)
Type of diabetes, n (%)						
Type 1 diabetes	15,728 (7.8)	9,490 (5.5)	4,567 (21.6)	1,131 (17.2)	227 (19.6)	313 (24.2)
Type 2 diabetes	152,622 (75.6)	139,696 (81.3)	9,494 (45.0)	2,678 (40.7)	418 (36.1)	336 (26.0)
Unknown	33,595 (16.6)	22,604 (13.2)	7,059 (33.4)	2,774 (42.1)	514 (44.3)	644 (49.8)
Duration of diabetes, years (IQR)						
Type 1 diabetes	15.4 (6.7;20.2)	9.7 (3.6;18.7)	19.6 (15.1;21.0)	19.7 (16.4;21.4)	19.5 (16.8;20.7)	20.2 (19.5;21.5)
Type 2 diabetes	5.3 (2.0;9.7)	5.0 (1.9;9.1)	10.5 (5.3;15.4)	11.0 (5.4;15.8)	11.2 (5.4;15.8)	12.9 (8.0;17.7)
Marital status, n (%)						
Never married	30,042 (14.9)	24,524 (14.3)	3,775 (17.9)	1,245 (18.9)	283 (24.4)	215 (16.6)
Married or living with someone	116,592 (57.7)	99,844 (58.1)	11,814 (55.9)	3,630 (55.1)	586 (50.6)	718 (55.5)
Widowed or divorced	55,311 (27.4)	47,422 (27.6)	5,531 (26.2)	1,708 (25.9)	290 (25.0)	360 (27.8)
Charlson Comorbidity Index, n (%)						
0 (low)	147,310 (72.9)	129,905 (75.6)	12,726 (60.3)	3,548 (53.9)	575 (49.6)	556 (43.0)
1 (moderate low)	25,987 (12.9)	18,245 (10.6)	5,005 (23.7)	1,889 (28.7)	386 (33.3)	462 (35.7)
2 (Moderate high)	18,153 (9.0)	15,436 (9.0)	1,851 (8.8)	614 (9.3)	114 (9.8)	138 (10.7)
≥3 (high)	10,495 (5.2)	8,204 (4.8)	1,538 (7.3)	532 (8.1)	84 (7.2)	137 (10.6)
Use of medication, n (%)						
Insulin	64,672 (32.0)	43,344 (25.2)	14,375 (68.1)	4,970 (75.5)	914 (78.9)	1,069 (82.7)
Antidiabetic drugs	154,069	136,384	12,469	3,988	664	564

(non-insulin)	(76.3)	(79.4)	(59.0)	(60.6)	(57.3)	(43.6)
Cholesterol lowering drugs	149,219 (73.9)	127,681 (74.3)	15,073 (71.4)	4,686 (71.2)	778 (67.1)	1,001 (77.4)
Anti-hypertensive drugs	150,825 (74.7)	127,955 (74.5)	15,853 (75.1)	5,063 (76.9)	859 (74.1)	1,095 (84.7)

Data are presented as percentage unless otherwise indicated. IQR = Interquartile range. PDR = proliferative diabetic retinopathy. DR = diabetic retinopathy. Antidiabetic drug treatment refers to all anti-diabetic drugs exclusive insulin. Adapted and modified from Paper I.

Table A2 (Paper II): Baseline differences between cases with subsequent progression to PDR and control patients with no or minimal progression in diabetic retinopathy (ICDR-level 0 or 1).

	Cases	Controls	Total
Individuals, n	39	89	128
Eyes, n	52	107	159
Baseline DR-level 0, n (%)	33 (63.5%)	63 (59.9%)	96 (60.4%)
Demographics			
Sex, n (%) female	21 (53.8%)	47 (52.2%)	68 (52.7%)
Age, years (IQR)	42.0 (24.0-60.0)	39.5 (25.0-60.0)	40.0 (25.0-60.0)
Diabetes type 1	25 (64.1%)	57 (63.3%)	82 (63.6%)
2	14 (35.9%)	33 (36.7%)	47 (36.4%)
Observation time, years (IQR)	10.8 (6.5-11.8)	10.4 (7.2-11.6)	10.5 (7.0-11.7)
Diabetes duration, years (IQR)	19.0 (12.0-29.0)	14.0 (10.0-21.0)	6.0 (2.0-13.0)
Para clinical measures			
Body Mass Index, kg/m^2	28.8 (23.9-33.8)	26.7 (24.2-30.1)	27.7 (24.0-31.4)
Systolic blood pressure, mmHg (IQR)	130 (120-140)	130 (118-137)	130 (120-140)
Diastolic blood pressure, mmHg (IQR)	80 (74-85)	75 (70-80)	77 (70-82)
Blood samples			
HbA1c, mmol/mol (IQR)	73 (60-90)	55 (46-65)	57 (50-75)
Serum-creatinine, µmol/l (IQR)	83.0 (76.0-92.0)	78.5 (66.0-91.0)	80.0 (68.5-91.5)
Low-density lipoprotein, mmol/l (IQR)	2.45 (2.10-2.90)	2.36 (1.90-2.80)	2.40 (1.90-2.80)
High-density lipoprotein, mmol/l (IQR)	1.49 (1.13-2.00)	1.35 (1.08-1.87)	1.39 (1.09-1.88)
Triglyceride, mmol/l (IQR)	1.32 (1.03-2.40)	1.16 (0.76-1.69)	1.21 (0.90-1.93)
Medications			
Insulin, n (%) yes	34 (87.2%)	70 (77.8%)	104 (80.6%)
Retinal vascular geometry			
CRAE, µm (IQR)	229 (211-250)	227 (210-247)	228 (211-248)
CRVE, µm (IQR)	289 (272-318)	290 (268-312)	290 (270-315)
Arterio-to-venule ratio	0.78 (0.74-0.85)	0.77 (0.72-0.85)	0.77 (0.72-0.85)

Continuous variables are medians (IQR), categorical as numbers (%). Case-control differences are individual-level, except CRAE, CRVE, and AVR (eye-level). Observation time spans first to last screening or PDR diagnosis; diabetes duration is from onset to last screening. Adapted and reprinted from Paper II.

AMERICAN ACADEMY OF OPHTHALMOLOGY®

Five-Year Incidence of Proliferative Diabetic Retinopathy and Associated Risk Factors in a Nationwide Cohort of 201 945 Danish Patients with Diabetes

Sebastian Dinesen, MD,[1,2,3] *Lonny Stokholm, PhD,*[4] *Yousif Subhi, PhD,*[2,5] *Tunde Peto, PhD,*[6] *Thiusius Rajeeth Savarimuthu, PhD,*[7] *Nis Andersen, PhD,*[8] *Jens Andresen, PhD,*[8] *Toke Bek, PhD, DMSci,*[9] *Javad Hajari, PhD,*[5] *Steffen Heegaard, PhD, DMSci,*[5] *Kurt Højlund, PhD, DMSci,*[2,3] *Caroline Schmidt Laugesen, PhD,*[10] *Ryo Kawasaki, PhD,*[2,11] *Sören Möller, PhD,*[2,4] *Katja Schielke, PhD,*[12] *Anne Suhr Thykjær, MD,*[1,2] *Frederik Pedersen, MD,*[1,2] *Jakob Grauslund, PhD, DMSci*[1,2,3]

Purpose: To evaluate the proliferative diabetic retinopathy (PDR) progression rates and identify the demographic and clinical characteristics of patients who later developed PDR compared with patients who did not progress to that state.

Design: A national 5-year register-based cohort study including 201 945 patients with diabetes.

Subjects: Patients with diabetes who had attended the Danish national screening program (2013–2018) for diabetic retinopathy (DR).

Methods: We used the first screening episode as the index date and included both eyes of patients with and without subsequent progression of PDR. Data were linked with various national health registries to investigate relevant clinical and demographic parameters. The International Clinical Retinopathy Disease Scale was used to classify DR, with no DR as level 0, mild DR as level 1, moderate DR as level 2, severe DR as level 3, and PDR as level 4.

Main Outcome Measures: Hazard ratios (HRs) for incident PDR for all relevant demographic and clinical parameters and 1-, 3-, and 5-year incidence rates of PDR according to baseline DR level.

Results: Progression to PDR within 5 years was identified in 2384 eyes of 1780 patients. Proliferative diabetic retinopathy progression rates from baseline DR level 3 at 1, 3 and 5 years were 3.6%, 10.9%, and 14.7%, respectively. The median number of visits was 3 (interquartile range, 1–4). Progression to PDR was predicted in a multivariable model by duration of diabetes (HR, 4.66 per 10 years; 95% confidence interval [CI], 4.05–5.37), type 1 diabetes (HR, 9.61; 95% CI, 8.01–11.53), a Charlson Comorbidity Index score of > 0 (score 1: HR, 4.62; 95% CI, 4.14–5.15; score 2: HR, 2.28; 95% CI, 1.90–2.74; score ≥ 3: HR, 4.28; 95% CI, 3.54–5.17), use of insulin (HR, 5.33; 95% CI, 4.49–6.33), and use of antihypertensive medications (HR, 2.23; 95% CI, 1.90–2.61).

Conclusions: In a 5-year longitudinal study of an entire screening nation, we found increased risk of PDR with increasing baseline DR levels, longer duration of diabetes, type 1 diabetes, systemic comorbidity, use of insulin, and blood pressure–lowering medications. Most interestingly, we found lower risk of progression from DR level 3 to PDR compared with that in previous studies.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2023;3:100291 © 2023 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Diabetic retinopathy (DR) is the most common long-term complication of diabetes mellitus[1,2] and the most common cause of vision loss in the adult population.[1,3] We offer systematic DR screening to reduce the risk of vision impairment and blindness as a free, tax-funded service in Denmark performed according to evidence-based national guidelines.[4] Trained ophthalmologists perform the grading, and 2-field or more fundus photography is the standard of care.

Systematic reviews conducted to clarify the risk factors for development of DR and incident proliferative DR (PDR) showed that a longer duration of diabetes, higher levels of blood glucose, type 1 diabetes, and hypertension are most commonly associated.[5,6]

Few prospective population-based studies have investigated the incidence of and associated risk factors for progression to PDR in both type 1 and type 2 diabetes.[7–13]

The study group investigated the existing incidence and risk factors for progression to PDR according to baseline characteristics in a national 5-year cohort of Danish patients with type 1 or type 2 diabetes. Previous studies that included both

https://doi.org/10.1016/j.xops.2023.100291
ISSN 2666-9145/23

patients with type 1 and those with type 2 diabetes were either of older date,[11,12] investigated relatively small populations,[7,9] or were focused on specific ethnic populations.[7,10]

Our aim was to evaluate the rate of progression to PDR and identify the demographic and clinical characteristics of patients with subsequent progression to PDR compared with patients who did not progress to that state. We based the study on data from the Danish Registry of Diabetic Retinopathy (DiaBase) between 2013 and 2018.

Methods

This was a cohort study based on data from Danish national registers that investigated eyes from patients with and without progression to PDR. Data from all patients with diabetes aged > 18 years who attend the Danish national screening program for DR are collected in DiaBase, which is a national clinical-quality database.[14] It contains information on the level of DR, visual acuity, date for screenings, indication for screening, and recommended time interval to next eye screening. Selected hospital departments and practicing opthalmologists offer the screening service performed in accordance with Danish guidelines for DR.[4] We included data on DR level and dates for screening.

The Danish National Patient Register includes all patient contacts within all Danish hospital facilities. It contains codes in the form of International Classification of Diseases 10 and Related Health Problems codes,[15] operations codes, treatment codes, date of admission, hospital codes, and geographical information; of these, we included data on diagnostic and treatment codes along with the date of admission.

The Danish National Prescription Registry holds information on medical prescriptions coded according to Anatomical Therapeutic Chemical Classification System codes.[16] We included Anatomical Therapeutic Chemical Classification System codes and dates for medical prescriptions dispensed at Danish pharmacies.

The Danish Civil Registration System[17] includes the unique identification numbers assigned to all individuals living in Denmark, which allows for the linking of data between registers. From this register, we extracted data on sex, age, marital status, death, and migration.

Our study population (Figure 1) included all Danish patients with diabetes aged > 18 years who attended the Danish screening program. The patients needed a registration in DiaBase with their first visit between January 2, 2013, and December 30, 2018. We allowed the inclusion of data from both eyes if eligible; however, these data are not shown in Tables 1 and 2, in which the patients were represented by the level of DR in their worst eye.

We used the first screening episode as the index date, and each patient contributed, from the first screening episode to their last episode of screening, reach of the end point (PDR), death, or emigration, which could be at any time between January 2, 2013, and December 30, 2018.

The distribution of diabetes types in the population was determined by combining diagnostic codes (International Classification of Diseases 10 codes) for diabetes and codes for prescription (Anatomical Therapeutic Chemical Classification System codes) of insulin. Pedersen et al[18] published supplementary material with a detailed description of the diabetes distribution.

The definition of progression to PDR (level 4) was set as a progression to DR level 4 in DiaBase from any other DR level (levels 0–3) at baseline according to the International Clinical Diabetic Retinopathy disease severity scale.

Increased severity is indicated with higher numbers, with no DR indicating level 0, mild nonproliferative DR (NPDR) indicating level 1 (microaneurysms and/or dot hemorrhages only), moderate NPDR indicating level 2 (more than just microaneurysms and/or dot hemorrhages but less than severe NPDR), severe NPDR indicating level 3 (> 20 intraretinal hemorrhages in each of 4 quadrants or prominent intraretinal microvascular abnormalities in ≥ 1 quadrant and no PDR), and PDR indicating level 4 (neovascularization [active or treated] or vitreous/preretinal hemorrhage).[19]

We measured progression in both eyes separately. Consequently, some patients contribute with 1 eye and some with 2 eyes. We excluded eyes with the presence of PDR at baseline, with subsequent regression from PDR, and if DR level assessment was absent at all screening visits.

We included a number of covariates that originated from the aforementioned databases and registers. They were measured at the index date and comprised sex (female or male), age at the index date (in years), duration of diabetes (in years), type of diabetes (type 1, type 2, or unknown), marital status (never married or married), patient comorbidity according to Charlson Comorbidity Index score[20] (0 = low, 1 = moderate low, 2 = moderate high, or ≥ 3 = high), and use of medication (insulin, noninsulin glucose-lowering treatment, antihypertensives, and cholesterol-lowering medications).

We compared the clinical and demographic baseline characteristics (Table 1) of patients who had subsequent progression to PDR (cases) with those of patients who did not have progression, defined as International Clinical Retinopathy Disease Severity Scale level 0 to 3 (controls). Descriptive data are presented as medians with interquartile ranges for continuous variables and counts with proportions for categorical variables.

We used the k-sample test for the equality of medians and chi-square test for continuous and categorical variables to test for differences between groups (Table 2).

The principal end points were defined as the adjusted hazard ratios (HRs) of markers of incident PDR according to baseline differences between patients with and without progression to PDR and 1-, 3-, and 5-year incidence rates of PDR according to the level of DR at baseline.

We displayed the HR for subsequent progression to PDR in Figure 2, estimated for all independent variables using the Cox proportional hazard model. We fitted a crude model as a semiadjusted (sex and age) and a multivariable model adjusted for sex, age, civil status, type of diabetes, blood pressure-lowering medications, cholesterol-lowering medications, and a modified Charlson Comorbidity Index score (excluding diabetes). Patients entered the analyses at the index date and were observed until progression to PDR, death, emigration, or their last screening episode, whichever occurred first. We accounted for the presence of patients with progression in 2 eyes using the robust standard errors.

We modeled a competing risk analysis (Fine-Gray subdistribution hazard methods) with death as a competing cause of progression to PDR among patients registered with DR level 3 at baseline compared with patients with DR level 2 at baseline. This was consequently done because of the findings in the study.

We used Stata 17 (StataCorp) for statistical analyses, with P values of < 0.05 and 95% confidence intervals (CIs) that did not include 1.0 considered statistically significant.

The study is a part of the Ocular and Systemic complications in DR project that originates from the Danish Excellence Centre in Ophthalmic Epidemiology Study.[21] We performed the study in accordance with the preconditions of the Declaration of Helsinki. We obtained all relevant permissions from the Region of Southern Denmark's record of data processing activities (Journal nr. 18/61231) and the Danish Clinical Registries (DIABASE-2018-12-11). The requirement of individual consent from each patient was not applicable.

Table 1. Baseline Characteristics of Patients in the Danish Registry of Diabetic Retinopathy According to the Level of Diabetic Retinopathy in the Worst Eye

	Overall	Level 0	Level 1	Level 2	Level 3	Level 4
Number of patients, n	201 945	171 790	21 120	6583	1159	1293
Male sex, n (%)	114 107 (56.5)	95 842 (55.8)	12 517 (59.3)	4156 (63.1)	786 (67.8)	806 (62.3)
Age (y), median (IQR)	65.8 (55.6–73.2)	66.2 (56.3–73.4)	63.5 (51.5–72.2)	62.0 (51.4–70.7)	56.0 (45.6–66.0)	64.1 (52.2–72.0)
Type of diabetes, n (%)						
Type 1 diabetes	15 728 (7.8)	9490 (5.5)	4567 (21.6)	1131 (17.2)	227 (19.6)	313 (24.2)
Type 2 diabetes	152 622 (75.6)	139 696 (81.3)	9494 (45.0)	2678 (40.7)	418 (36.1)	336 (26.0)
Unknown	33 595 (16.6)	22 604 (13.2)	7059 (33.4)	2774 (42.1)	514 (44.3)	644 (49.8)
Duration of diabetes (y), median (IQR)						
Type 1 diabetes	15.4 (6.7–20.2)	9.7 (3.6–18.7)	19.6 (15.1–21.0)	19.7 (16.4–21.4)	19.5 (16.8–20.7)	20.2 (19.5–21.5)
Type 2 diabetes	5.3 (2.0–9.7)	5.0 (1.9–9.1)	10.5 (5.3–15.4)	11.0 (5.4–15.8)	11.2 (5.4–15.8)	12.9 (8.0–17.7)
Marital status, n (%)						
Never married	30 042 (14.9)	24 524 (14.3)	3775 (17.9)	1245 (18.9)	283 (24.4)	215 (16.6)
Married or living with someone	116 592 (57.7)	99 844 (58.1)	11 814 (55.9)	3630 (55.1)	586 (50.6)	718 (55.5)
Widowed or divorced	55 311 (27.4)	47 422 (27.6)	5531 (26.2)	1708 (25.9)	290 (25.0)	360 (27.8)
Charlson Comorbidity Index, n (%)						
0 (low)	147 310 (72.9)	129 905 (75.6)	12 726 (60.3)	3548 (53.9)	575 (49.6)	556 (43.0)
1 (moderate low)	25 987 (12.9)	18 245 (10.6)	5005 (23.7)	1889 (28.7)	386 (33.3)	462 (35.7)
2 (Moderate high)	18 153 (9.0)	15 436 (9.0)	1851 (8.8)	614 (9.3)	114 (9.8)	138 (10.7)
≥ 3 (high)	10 495 (5.2)	8204 (4.8)	1538 (7.3)	532 (8.1)	84 (7.2)	137 (10.6)
Use of medication, n (%)						
Insulin	64 672 (32.0)	43 344 (25.2)	14 375 (68.1)	4970 (75.5)	914 (78.9)	1069 (82.7)
Antidiabetic drugs (noninsulin)*	154 069 (76.3)	136 384 (79.4)	12 469 (59.0)	3988 (60.6)	664 (57.3)	564 (43.6)
Cholesterol-lowering drugs	149 219 (73.9)	127 681 (74.3)	15 073 (71.4)	4686 (71.2)	778 (67.1)	1001 (77.4)
Antihypertensive drugs	150 825 (74.7)	127 955 (74.5)	15 853 (75.1)	5063 (76.9)	859 (74.1)	1095 (84.7)

IQR = interquartile range.
*Antidiabetic drug treatment refers to all antidiabetic drugs exclusive insulin.

Table 2. Baseline Differences between Patients with and without Subsequent Progression to PDR in ≥ 1 Eye in the Danish Registry of Diabetic Retinopathy

	All	Subsequent Progression to PDR	No Subsequent Progression to PDR	*P* Value
Number of patients, n	201 945	1780	200 165	
Male sex, n (%)	114 107 (56.5)	1074 (60.3)	113 033 (56.5)	0.001
Age (y), median (IQR)	65.8 (55.6–73.2)	59.6 (48.0–68.9)	65.8 (55.6–73.2)	0.001
Type of diabetes, n (%)				< 0.001
Type 1 diabetes	15 728 (7.8)	511 (28.7)	15 217 (7.6)	
Type 2 diabetes	152 622 (75.6)	421 (23.7)	152 201 (76.0)	
Unknown	33 595 (16.6)	848 (47.6)	32 747(16.4)	
Duration of diabetes (y), median (IQR)				< 0.001
Type 1 diabetes	15.4 (6.7–20.2)	19.9 (19.1–21.3)	14.9 (6.5–20.2)	
Type 2 diabetes	5.3 (2.0–9.7)	11.7 (5.8–17.4)	5.3 (2.0–9.7)	
Marital status, n (%)				< 0.001
Never married	30 042 (14.9)	390 (21.9)	29 652 (14.8)	
Married	116 592 (57.7)	1005 (56.5)	115 587 (57.7)	
Widowed or divorced	55 311 (27.4)	385 (21.6)	54 926 (27.4)	
Charlson Comorbidity Index score, n (%)				< 0.001
0 (low)	147 310 (72.9)	770 (43.3)	146 540 (73.2)	
1 (moderate low)	25 987 (12.9)	698 (39.2)	25 289 (12.6)	
2 (Moderate high)	18 153 (9.0)	163 (9.2)	17 990 (9.0)	
≥ 3 (high)	10 495 (5.2)	149 (8.4)	10 346 (5.2)	
Use of medication, n (%)				
Insulin	64 672 (32.0)	1536 (86.3)	63 136 (31.5)	< 0.001
Antidiabetic drugs (noninsulin)	154 069 (76.3)	759 (42.6)	153 310 (76.6)	< 0.001
Antihypertensive drugs	150 825 (74.7)	1424 (80.0)	149 401 (74.6)	< 0.001
Cholesterol-lowering drugs	149 219 (73.9)	1294 (72.7)	147 925 (73.9)	0.29

IQR = interquartile range; PDR = proliferative diabetic retinopathy.

Results

We found that 201 945 patients were eligible for inclusion, of which we identified 1780 (0.9%) patients and 2384 (0.6%) eyes with progression to PDR in 5 years of follow-up. There were 510, 1041, 535, and 298 eyes with progression from baseline DR levels 0, 1, 2, and 3, respectively, which equaled to rates of 0.1%, 2.7%, 4.7%, and 14.7%, respectively. The median number of visits was 3 (interquartile range, 1–4).

The population (Table 1) consisted predominantly of men (56.5%) and elderly individuals (median age, 65.8 years). Type 2 diabetes (75.6%) was more prevalent than type 1 diabetes (7.8%), and the last proportion (16.6%) had an unknown type of diabetes. The duration of diabetes at baseline was longer in patients with type 1 diabetes (median, 15.4 vs. 5.3 years) than in those with type 2 diabetes. The duration of diabetes increased with higher DR levels at baseline for type 2 diabetes (5.0 vs. 10.5 vs. 11.0 vs. 11.2 vs. 12.9 years), but this was not consistent for type 1 diabetes (9.7 vs. 19.6 vs. 19.7 vs. 19.5 vs. 20.2 years). Patients with higher DR levels at baseline were more likely to have never been married (14.3% vs. 17.9% vs. 18.9% vs. 24.4% vs. 16.6%). Increasing DR levels (4.8% vs. 7.3% vs. 8.1% vs. 7.2% vs. 10.6%) were associated with higher Charlson Comorbidity Index scores (≥ 3). Insulin treatment was more prevalent (25.2% vs. 68.1% vs. 75.5% vs. 78.9% vs. 82.7%), whereas use of noninsulin glucose-lowering medications (79.4% vs. 59.0% vs. 60.6% vs. 57.3% vs. 43.6%) was lower with increasing levels of DR. There was no consistent decrease or increase in the use of antihypertensives and cholesterol-lowering drugs in relation to increasing DR levels.

Baseline differences between patients with and without subsequent progression to PDR (Table 2) showed that male sex (60.3 vs. 56.5%, $P = 0.001$) and lower age (59.6 vs. 65.8 years, $P = 0.001$) were associated with incident PDR. Furthermore, progression to PDR was associated with longer duration of diabetes (type 1 diabetes: 19.9 vs. 14.9 years, $P < 0.001$; type 2 diabetes: 11.7 vs. 5.3 years, $P < 0.001$), type 1 diabetes (28.7% vs. 7.6%, $P < 0.001$), never being married (21.9 vs. 14.8%, $P < 0.001$), and use of insulin (86.3 vs. 31.5%, $P < 0.001$) and blood pressure–lowering medication (80.0 vs. 74.6%, $P < 0.001$).

We compared patients without PDR progression with patients with subsequent progression to PDR in a Cox proportional hazard model, and the latter (Figure 2) had a longer duration of diabetes at baseline (HR, 4.66 per 10 years; 95% CI, 4.05–5.37) and a Charlson Comorbidity Index score of > 0 (score 1: HR, 4.62; 95% CI, 4.14–5.15; score 2: HR, 2.28; 95% CI, 1.90–2.74; score ≥ 3: HR, 4.28; 95% CI, 3.54–5.17) and were more likely to never have been married (HR, 1.36; 95% CI, 1.19–1.56), to have type 1 diabetes (HR, 9.61; 95% CI, 8.01–11.53), to receive insulin (HR, 5.33; 95% CI,

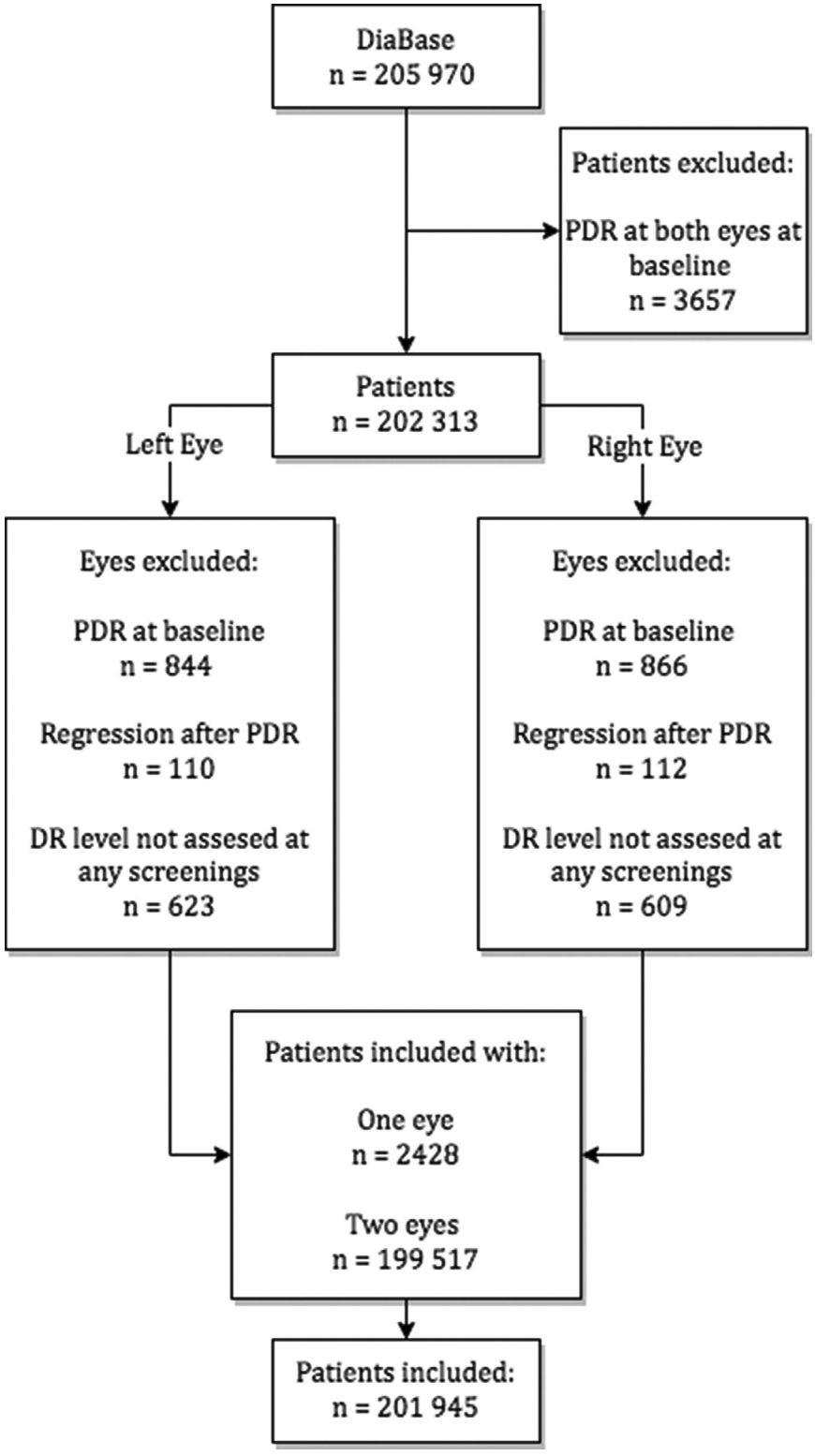


Figure 1. A flowchart displaying the selected study population. DiaBase = The Danish Registry of Diabetic Retinopathy; DR = diabetic retinopathy; PDR = proliferative diabetic retinopathy.

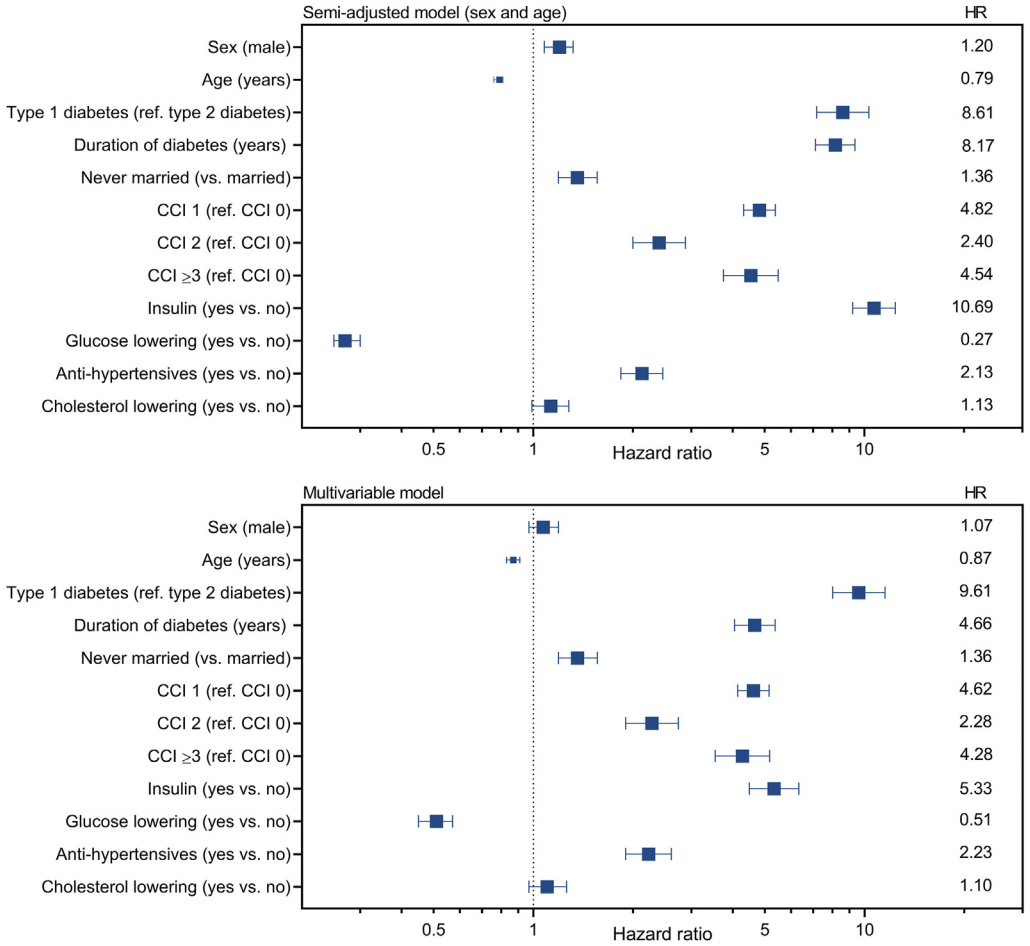


Figure 2. Forrest plots of the hazard ratio for subsequent progression to proliferative diabetic retinopathy according to baseline characteristics. The figure shows a semiadjusted (sex and age) and a multivariable Cox proportional hazard model. Cholesterol lowering refers to all types of cholesterol-lowering medications. Glucose lowering refers to glucose-lowering medications exclusive of insulins. CCI = Charlson Comorbidity Index; HR = hazard ratio; ref. = reference.

4.49−6.33), and to be treated with blood pressure−lowering medication (HR, 2.23; 95% CI, 1.90−2.61).

There was a decreased risk with age at first eye screening per 10-year increment (HR, 0.87; 95% CI, 0.83−0.91) and noninsulin glucose-lowering treatment (HR, 0.51; 95% CI, 0.45−0.57) but no differences in risk regarding cholesterol-lowering drugs (HR, 1.10; 95% CI, 0.97−1.26) and sex (HR, 1.07; 95% CI 0.97−1.19).

Patients with DR level 3 at baseline had an increased risk of progression to PDR when compared with that in those with DR level 2 (HR, 3.25; 95% CI, 2.74−3.87), and when death was examined as a competing cause to PDR, the findings indicated no effect of death on the risk of progression (HR, 3.25; 95% CI, 2.74−3.87).

We performed temporal analysis (Table 3), which showed the incidence rates of PDR within 1, 3, and 5 years according to the patients' baseline DR level. One-, 3- and 5-year incidence rates were 0.6, 0.6, and 0.7 events per 1000 person-years for patients with DR level 0; 8.3, 10.8, and 11.6 events per 1000 person-years for patients with DR level 1; 13.9, 17.9, and 19.9 events per 1000 person-years for patients with DR level 2; and 52.7, 67.9, and 72.7 events per 1000 person-years for patients with DR level 3, respectively.

A total of 4237 (2.1%) patients with DR levels 0 to 3 at baseline had VEGF inhibitor injections at any time from 2007 to 2018. The treatment was distributed as 2260 (1.3%), 1100 (5.5%), 631 (10.6%), and 246 (26.9%) patients with DR levels 0, 1, 2, and 3, respectively, at baseline.

Discussion

This was a national 5-year prospective cohort study of 201 945 patients with type 1 or type 2 diabetes who attended the national screening program for DR. We identified type 1 diabetes, duration of diabetes, and insulin treatment as the most prevalent risk factors for subsequent progression to PDR.

We found that 0.9% of the cohort progressed to PDR over a 5-year period. A meta-analysis, performed by Wong et al,[22] included 27 120 patients with diabetes (48% with type

Table 3. Risk of Progression to Proliferative DR within 1, 3, and 5 Years According to the Level of DR at the Time of the First Registration in the Danish Registry of Diabetic Retinopathy*

		1 Year			3 Years			5 Years		
Level of DR	**Eyes at Risk**	*Events*	*Observation Time*	*Incidence Rate per 1000 person-years*	*Events*	*Observation Time*	*Incidence Rate per 1000 person-years*	*Events*	*Observation Time*	*Incidence Rate per 1000 person-years*
All	401 462	570	282 631	2.0	1759	680 700	2.6	2384	817 743	2.9
0	352 143	138	243 885	0.6	382	584 546	0.6	510	696 917	0.7
1	36 280	239	28 800	8.3	778	71 903	10.8	1041	89 823	11.6
2	11 006	119	8542	13.9	377	20 979	17.9	535	26 900	19.9
3	2033	74	1403	52.7	222	3270	67.9	298	4101	72.7

DR = diabetic retinopathy.
*A total of 2384 eyes had progression to proliferative DR over 5 years. The individuals at risk were 201 945.

2 diabetes) and reported a 5-year risk of progression to PDR of 6.4% on the basis of studies from 1986 to 2008. Thus, they reported substantially higher rates of progression to PDR than those in our study. This study contributes new knowledge about the rate of progression in well-controlled patients with diabetes. The study was performed nationwide and consisted of both patients with type 1 diabetes and patients with type 2 diabetes, of which a large proportion were relatively healthy patients with type 2 diabetes. Older studies consisted of poorly regulated high-risk populations with diabetes, as in the meta-analysis by Wong et al,[22] among which, some studies solely investigated patients with type 1 diabetes. Our results indicate that good systemic control helps with reducing diabetic eye complications.

It was difficult to compare our study with others because of differences in the follow-up time, method, and study population. A prospective cohort study from 2007 to 2014 by Romero-Aroca et al[23] of 15 396 patients with type 1 (6.6%) and type 2 (93.4%) diabetes reported 19 PDR progression patient cases (0.1%) from any baseline DR level in 8 years of follow-up time. In comparison, we reported 1780 patients (0.9%). They included 6.6% of patients with type 1 diabetes compared with 9.5% in our study.

A 5-year (2005−2009) prospective study from Scotland[24] reported an overall PDR progression rate of 0.9%, solely investigating 49 763 patients with type 2 diabetes, whereas a prospective study from England[25] reported a 17-year (1990−2006) PDR progression rate of 0.9% from any baseline DR level, including both type 1 (1%) and type 2 diabetes (99%).

The current study found lower 1-, 3-, and 5-year PDR progression rates for individuals with baseline DR level 3 (severe NPDR) compared with those in earlier reports.[26] The ETDRS Research Group, which laid the foundation for the development of the International Clinical Retinopathy Disease Severity classification scale,[19] found 1-, 3-, and 5-year risks of progression to high-risk PDR to be 17.1%, 44.4%, and 57.8%, respectively. We found the 1-, 3- and 5-year risks to be 3.6%, 10.9%, and 14.7%, respectively. Although the current recommendations[4] on screening intervals for patients with DR level 3 are based on much higher progression rates,[26−28] our study indicates that screening intervals can be reconsidered if all systemic factors (blood pressure and hemoglobin A1c) are known and can be taken into account. Our results strengthen the notion that clinicians should control systemic factors appropriately (hemoglobin A1c and blood pressure), treat diabetic macular edema, and care for patients in need of special care (pregnancy, bariatric surgery, and mental or physical vulnerability) when proposing individual screening intervals.[4]

We performed a competing risk analysis with mortality as a competing outcome for individuals with DR level 3 (compared with DR level 2) to assess whether the low incidence rates were due to death occurring before the individuals had progression to PDR. The HR from the competing risk analysis was compared with the HR for progression to PDR between DR levels 2 and 3 and could not explain the low incidence rates. These findings add further substance to the trend that PDR progression rates are declining compared with those in previous decades in populations with relatively good control and well-characterized and cared for risk factors.[22,29] This seems particularly true for those with DR level 3 at baseline, which gives a strong argument for better prevention, monitoring, and treatment of diabetes to reduce the risk of PDR.

The risk of progression to PDR within 1, 3, and 5 years was found to increase with higher DR levels at baseline (Table 3). The findings from a population-based 5-year cohort study by Janghorbani et al[8] align with our findings showing that baseline DR level is the main risk factor for incident PDR regardless of diabetes type. They also found that insulin-treated diabetes compared with noninsulin-treated diabetes, duration of diabetes, and poor glycemic control were all risk factors for subsequent progression to PDR, with the first 2 being statistically significant risk factors in this study as well.

Although it is well known that duration of diabetes is a strong risk factor for DR and that every patient with onset of diabetes before the age of 15 years will develop PDR,[13] we were encouraged by the observation that duration of diabetes as a risk factor was mostly driven by the larger difference in patients with type 2 diabetes (Table 2).

The study represents an entire national screening population, which is considered a major strength; however, we need to address the limitations. First, the reports of DR grading in DiaBase rely on data input from many different

sources. To address this, the accuracy has recently been validated by Thykjær et al.[30] In 458 randomly chosen eyes, there was a full 5-step DR grading agreement of 93% between ophthalmologists reporting to DiaBase and a certified expert grader, corresponding to an overall agreement of 96% (κ, 0.89) and 90% (κ, 0.76) for practicing ophthalmologists and hospital-based grading centers, respectively. Likewise, a virtual ocular learning platform has been nationally launched to support training in DR grading of Danish ophthalmologists.[31]

Second, data on glycaemic control, hypertension, body mass index, smoking, and physical activity were not included because of unavailability. Third, it was not possible to include data on diabetic macular edema because of inconsistent reporting in DiaBase. As a sensitivity analysis, we used intravitreal treatment with VEGF inhibitors as a proxy marker for diabetic macular edema. Only 2.1% of patients received intravitreal injections between 2007 and 2018, with the highest number administered to patients with severe NPDR at baseline, of which 26.9% subsequently received intravitreal treatment. Hence, we cannot exclude the possibility of a potential causal pathway between diabetic macular edema and PDR for a minority of patients.

This 5-year longitudinal study of a national cohort consisted of patients with diabetes attending the screening program for DR. We identified, with the first screening episode as the baseline, that patients with subsequent progression to PDR differed from patients with no or minimal progression. Duration of diabetes, type of diabetes, and use of insulin were all identified as the most important risk factors for incident PDR. Finally, the incidence rates of progression to PDR were lower than those reported previously and in particular for patients with DR level 3 at baseline.

Footnotes and Disclosures

Originally received: June 17, 2022.
Final revision: February 16, 2023.
Accepted: February 17, 2023.
Available online: February 24, 2023. Manuscript no. XOPS-D-22-00129R2.

[1] Department of Ophthalmology, Odense University Hospital, Odense, Denmark.

[2] Department of Clinical Research, University of Southern Denmark, Odense, Denmark.

[3] Steno Diabetes Center Odense, Odense University Hospital, Odense, Denmark.

[4] Open Patient Data Explorative Network, Department of Clinical Research, University of Southern Denmark and Odense University Hospital, Odense, Denmark.

[5] Department of Ophthalmology, Rigshospitalet, Glostrup, Denmark.

[6] Centre for Public Health, Queen's University Belfast, Belfast, United Kingdom.

[7] The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark.

[8] Organization of Danish Practicing Ophthalmologists, Copenhagen, Denmark.

[9] Department of Ophthalmology, Aarhus University Hospital, Aarhus, Denmark.

[10] Department of Ophthalmology, Zealand University Hospital Roskilde, Roskilde, Denmark.

[11] Department of Vision Informatics, University of Osaka, Osaka, Japan.

[12] Department of Ophthalmology, Aalborg University Hospital, Aalborg, Denmark.

Disclosures:

All authors have completed and submitted the ICMJE disclosure form.

The authors have made the following disclosures: R.K.: Grant – Topcon, Senju, Novartis, Nanolux, Tamron, and WHILL, outside the submitted work; Personal fees – Novartis, Bayer, Office Future, Nanolux, Kowa, Santen, and Bausch Lomb, outside the submitted work.

J.G.: Grant – VELUX FONDEN; Personal fees – Bayer, Novartis, Roche, Allergan, and Santen, outside the submitted work.

Supported by Steno Diabetes Center Odense, Odense, Denmark, and the Region of Southern Denmark, Denmark. The funding organizations had no role in the design or conduct of this research.

Preliminary data presented at The European Association for the Study of Diabetes (EASD) Eye Complications 31st Conference. Presented at the Eye Complication Study Groups Annual Meeting, October 30, 2021.

HUMAN SUBJECTS: Human subjects were included in this study. The study was performed in accordance with the tenets of the Declaration of Helsinki. All relevant permissions were obtained from the Region of Southern Denmark's record of data processing activities (Journal nr. 18/61231) and the Danish Clinical Registries (DIABASE-2018-12-11). The requirement of individual consent from each patient was not applicable.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Dinesen, Stokholm, Andersen, Andresen, Bek, Hajari, Heegaard, Højlund, Laugesen, Kawasaki, Möller, Schielke, Thykjær, Grauslund

Data collection: Dinesen, Stokholm, Andersen, Højlund, Grauslund

Analysis and interpretation: Dinesen, Stokholm, Peto, Grauslund

Obtained funding: Study was performed as part of regular employment duties at the Department of Ophthalmology, Odense, Denmark.

Overall responsibility: Dinesen, Stokholm, Subhi, Peto, Savarimuthu, Andersen, Adresen, Bek, Hajari, Heegaard, Højlund, Laugesen, Kawasaki, Möller, Schielke, Thykjær, Pedersen, Grauslund

Abbreviations and Acronyms:
CI = confidence interval; **DiaBase** = The Danish Registry of Diabetic Retinopathy; **DR** = diabetic retinopathy; **HR** = hazard ratio; **NPDR** = nonproliferative diabetic retinopathy; **PDR** = proliferative diabetic retinopathy.

Keywords:
Epidemiology, Proliferative diabetic retinopathy, risk factor.

Correspondence:
Sebastian Dinesen, MD, Department of Ophthalmology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense, Denmark. E-mail: sebastian.dinesen@rsyd.dk.

References

1. Carstensen B, Rønn PF, Jørgensen ME. Prevalence, incidence and mortality of type 1 and type 2 diabetes in Denmark 1996-2016. *BMJ Open Diabetes Res Care*. 2020;8:e001071.
2. Grauslund J, Green A, Kawasaki R, et al. Retinal vascular fractals and microvascular and macrovascular complications in type 1 diabetes. *Ophthalmology*. 2010;117:1400−1405.
3. Grauslund J, Green A, Sjølie AK. Blindness in a 25-year follow-up of a population-based cohort of Danish type 1 diabetic patients. *Ophthalmology*. 2009;116:2170−2174.
4. Grauslund J, Andersen N, Andresen J, et al. Evidence-based Danish guidelines for screening of diabetic retinopathy. *Acta Ophthalmol*. 2018;96:763−769.
5. Yau JW, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*. 2012;35:556−564.
6. Song P, Yu J, Chan KY, et al. Prevalence, risk factors and burden of diabetic retinopathy in China: a systematic review and meta-analysis. *J Glob Health*. 2018;8:010803.
7. Varma R, Choudhury F, Klein R, et al. Four-year incidence and progression of diabetic retinopathy and macular edema: the Los Angeles Latino Eye Study. *Am J Ophthalmol*. 2010;149:752−761.e3.
8. Janghorbani M, Jones RB, Allison SP. Incidence of and risk factors for proliferative retinopathy and its association with blindness among diabetes clinic attenders. *Ophthalmic Epidemiol*. 2000;7:225−241.
9. McCarty DJ, Fu CL, Harper CA, et al. Five-year incidence of diabetic retinopathy in the Melbourne Visual Impairment Project. *Clin Exp Ophthalmol*. 2003;31:397−402.
10. Leske MC, Wu SY, Hennis A, et al. Nine-year incidence of diabetic retinopathy in the Barbados Eye Studies. *Arch Ophthalmol*. 2006;124:250−255.
11. Klein R, Klein BE, Moss SE, et al. The Wisconsin Epidemiologic Study of Diabetic Retinopathy. IX. Four-year incidence and progression of diabetic retinopathy when age at diagnosis is less than 30 years. *Arch Ophthalmol*. 1989;107:237−243.
12. Klein R, Klein BE, Moss SE, et al. The Wisconsin Epidemiologic Study of Diabetic Retinopathy. X. Four-year incidence and progression of diabetic retinopathy when age at diagnosis is 30 years or more. *Arch Ophthalmol*. 1989;107:244−249.
13. Bek T. Systemic risk factors contribute differently to the development of proliferative diabetic retinopathy and clinically significant macular oedema. *Diabetologia*. 2020;63:2462−2470.
14. Andersen N, Hjortdal JØ, Schielke KC, et al. The Danish Registry of Diabetic Retinopathy. *Clin Epidemiol*. 2016;8:613−619.
15. World Health Organization. International classification of disease and related health problems, tenth revision (ICD 10). Available at: https://www.who.int/publications/m/item/official-who-icd-10-updates-2019-package; Accessed February 1, 2020.
16. Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. *Scand J Public Health*. 2011;39(7 suppl):38−41.
17. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur J Epidemiol*. 2014;29:541−549.
18. Pedersen FN, Stokholm L, Pouwer F, et al. Diabetic retinopathy predicts risk of Alzheimer's disease: a Danish registry-based nationwide cohort study. *J Alzheimers Dis*. 2022;86:451−460.
19. Wilkinson CP, Ferris III FL, Klein RE, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. 2003;110:1677−1682.
20. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. 2011;173:676−682.
21. Grauslund J, Stokholm L, Ohm Kyvik K, et al. Interactions between ocular and systemic disease using national register-based data in the Danish Excellence Centre in Ophthalmic Epidemiology (DECODE-EYE): study perspective. *Acta Ophthalmol*. 2020;98:573−578.
22. Wong TY, Mwamburi M, Klein R, et al. Rates of progression in diabetic retinopathy during different time periods: a systematic review and meta-analysis. *Diabetes Care*. 2009;32:2307−2313.
23. Romero-Aroca P, de la Riva-Fernandez S, Valls-Mateu A, et al. Changes observed in diabetic retinopathy: eight-year follow-up of a Spanish population. *Br J Ophthalmol*. 2016;100:1366−1371.
24. Thomas RL, Dunstan F, Luzio SD, 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.
25. Jones CD, Greenwood RH, Misra A, Bachmann MO. Incidence and progression of diabetic retinopathy during 17 years of a population-based screening program in England. *Diabetes Care*. 2012;35:592−596.
26. Fundus photographic risk factors for progression of diabetic retinopathy. ETDRS report number 12. Early Treatment Diabetic Retinopathy Study Research Group. *Ophthalmology*. 1991;98(5 Suppl):823−833.
27. Klein R, Klein BE, Moss SE, et al. 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;102:527−532.
28. Klein R, Klein BE, Moss SE, et al. 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;102:520−526.
29. Sabanayagam C, Banu R, Chee ML, et al. Incidence and progression of diabetic retinopathy: a systematic review. *Lancet Diabetes Endocrinol*. 2019;7:140−149.
30. Thykjær A, Stokholm L, Kawasaki R, et al. ARVO Annual Meeting Abstract 2022. Inter grader agreement in the Danish screening program for diabetic retinopathy. https://iovs.arvojournals.org/article.aspx?articleid=2779555. Accessed June, 2022.
31. Andersen JKH, Hubel MS, Savarimuthu TR, et al. A digital online platform for education and certification of diabetic retinopathy health care professionals in the Region of Southern Denmark. *Acta Ophthalmol*. 2022;100:589−595.

Received: 12 May 2023 | Accepted: 18 September 2023

DOI: 10.1111/aos.15780

Acta Ophthalmologica

ORIGINAL ARTICLE

Retinal main vessel calibers and systemic markers for long-term development of proliferative diabetic retinopathy

Sebastian Dinesen[1,2,3] | **Lonny Stokholm**[4] | **Yousif Subhi**[2,5] | **Jan Erik Henriksen**[2,3] | **Thiusius Rajeeth Savarimuthu**[6] | **Tunde Peto**[7] | **Jakob Grauslund**[1,2,3]

[1]Department of Ophthalmology, Odense University Hospital, Odense, Denmark

[2]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[3]Steno Diabetes Center Odense, Odense University Hospital, Odense, Denmark

[4]Open Patient Data Explorative Network, Department of Clinical Research, University of Southern Denmark and Odense University Hospital, Odense, Denmark

[5]Department of Ophthalmology, Rigshospitalet, Glostrup, Denmark

[6]The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

[7]Centre for Public Health, Queen's University Belfast, Belfast, UK

Correspondence
Sebastian Dinesen, Department of Ophthalmology, Odense University Hospital, Sdr. Boulevard 29, Odense 5000, Denmark.
Email: sebastian.dinesen@rsyd.dk

Funding information
Steno Diabetes Center Odense, Grant/Award Number: 4836; The Region of Southern Denmark, Grant/Award Number: 823

Abstract

Purpose: To evaluate if retinal vascular calibers and systemic risk factors in patients with no or minimal diabetic retinopathy (DR) can predict risk of long-term progression to proliferative diabetic retinopathy (PDR).

Methods: This was a matched case–control study of patients with diabetes having no or minimal DR at baseline with (cases) or without (controls) subsequent development of PDR. We collected six-field, 45-degree retinal images, demographic and clinical data from the Funen Diabetes Database.

Results: We included 52 eyes from 39 cases and 107 eyes from 89 controls matched on sex, age, type of diabetes, time from first to last screening episode and baseline DR level. Cases had higher HbA1c (73 vs. 55 mmoL/moL; $p<0.001$), triglycerides (1.32 vs. 1.16 mmoL/L; $p=0.02$) and longer duration of diabetes (19 vs. 14 years; $p=0.01$), but the groups did not differ in calibers of retinal arterioles (229 vs. 227 μm; $p=0.49$), venules (289 vs. 290 μm; $p=0.83$) or the arterio-to-venule ratio (0.78 vs. 0.77; $p=0.86$).In a multivariable logistic regression model with cluster robust standard error, HbA1c (OR 1.54 per 10 mmoL/moL; 95%-CI: 1.15–2.07; $p=0.004$), triglyceride (OR 1.39 per 1 mmoL/L; 95%-CI: 1.03–1.86; $p=0.03$) and duration of diabetes (OR 1.09 per year; 95%-CI: 1.03–1.16; $p=0.003$) were independent risk factors for PDR.

Conclusion: Retinal vascular calibers did not predict long-term development of PDR in contrast to well-established risk factors like HbA1c, triglyceride and duration of diabetes.

KEYWORDS
proliferative diabetic retinopathy, prediction, retinal vascular calibers, risk factor

1 | INTRODUCTION

Proliferative diabetic retinopathy (PDR) is a vision-threatening complication of diabetes and is among the leading causes of preventable blindness (Carstensen et al., 2020). Early identification of patients at risk of PDR would be important in order to address modifiable risk factors (Yau et al., 2012) such as haemoglobin-A1c (HbA1c), blood lipids and blood pressure. In the last two decades, structural analysis of the retinal vascularity (Cheung et al., 2015) has been suggested as new clinical predictive measures for incidence and progression of diabetic retinopathy (DR).

An objective measure of retinal vascular calibre has been developed along with improvements in imaging the retina allowing evaluation of vessel calibre in DR (Ikram et al., 2013) and other vascular diseases (Ikram, de Jong, Bos, et al., 2006; Ikram, de Jong, van Dijk, et al., 2006; Ikram, Witteman, et al., 2006; Wong et al., 2001). Retinal venular diameter has been shown to predict DR progression and incident PDR, while retinal arteriolar narrowing has been seen to predict DR onset, but not more advanced DR stages (Wong, 2011).

The purpose of this study was to investigate the association of retinal vascular calibers and systemic factors with long-term PDR development in a population with diabetes but no or minimal DR at baseline.

2 | METHODS

2.1 | Design

This was a matched case–control study that investigated retinal vascular calibers and systemic markers in eyes with no or minimal DR at baseline with (cases) or without (controls) subsequent development of PDR.

2.2 | Database

The Funen Diabetes Database (FDDB) is a regional database containing demographic and clinical data from the Danish DR screening program for patients with diabetes mellitus that resides on Funen in Denmark (Adelborg et al., 2020).

We collected essential demographic information such as age and sex as well as clinical data including International Clinical Diabetic Retinopathy (ICDR) level, type of diabetes, screening dates, body mass index, systolic and diastolic blood pressure. Additionally, we obtained blood test results for low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglyceride, serum-creatinine and HbA1c as well as medication prescriptions for insulin.

2.3 | Study population

We included patients of all ages with type 1 or 2 diabetes who attended the Danish DR screening program from 2003 to 2019 (Larsen et al., 2017).

The cases had to be registered in the FDDB with one or two eyes having ICDR-level 0 (No DR) or 1 (microaneurysms and/or dot haemorrhages only) and develop PDR during the observation period from 2003 to 2019. When these criteria were fulfilled for one or both eyes, the local image database was retrospectively checked for availability of the initial retinal images with ICDR-level 0 or 1 and the agreement between the screening dates/grades in the image database and FDDB was ascertained. The patients were excluded when initial fundus images were unavailable (Figure 1: flowchart).

The first screening episode registered served as baseline for the study population from which we matched cases with controls (Figure 2). Case patients were individually matched to control patients with the same baseline level of DR, who did not progress beyond ICDR-level 0 (baseline DR-level 0) or ICDR-level 1 (baseline DR-level 1). Further matching included age, sex, type of diabetes and time from first to last screening (observation time). The study was designed for each case to be matched with three controls, but we only managed to match 11 of the case patients to a third suitable control patient that fulfilled the criteria for matching.

2.4 | Assessment of DR

All mydriatic digital fundus photos were obtained as six field, 45-degree images with Topcon TRC-50X (Topcon).

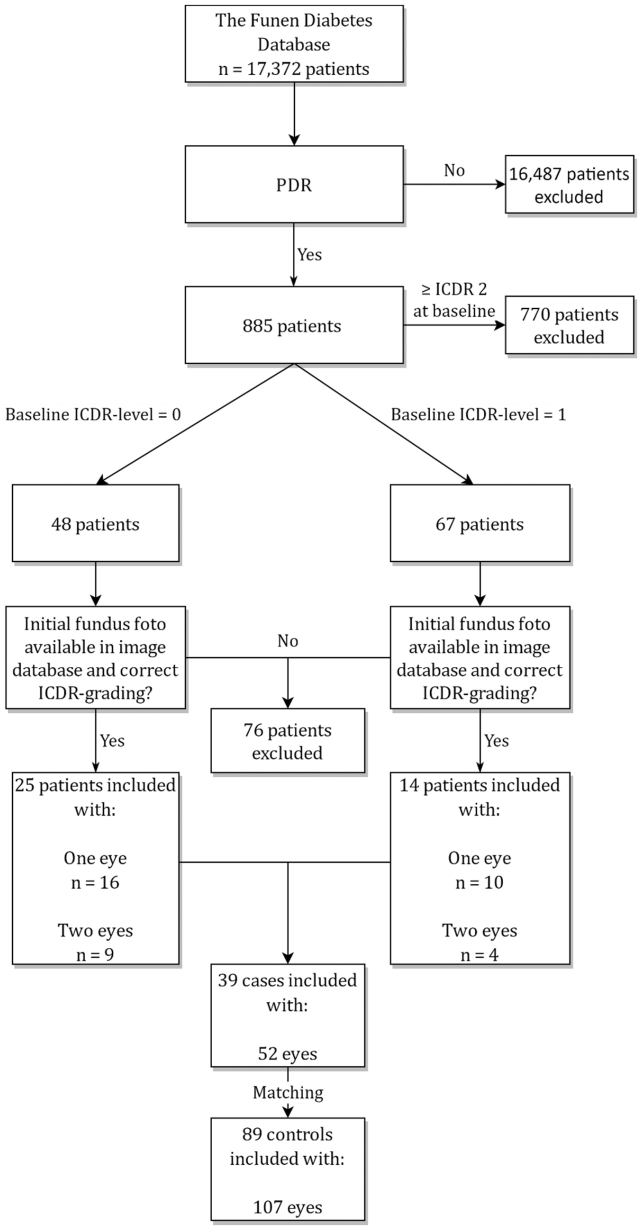


FIGURE 1 Flowchart for identification of case patients from the Funen Diabetes Database. ICDR, International Clinical Diabetic Retinopathy scale; PDR, proliferative diabetic retinopathy.

All fundus images were re-assessed for correct ICDR-grading by a certified grader according to the ICDR guidelines (Wilkinson et al., 2003) before enrolment in the study.

2.5 | Retinal vascular calibre analysis

The semi-automatic software VAMPIRE (version 3.2; Vessel Assessment and Measurement Platform for Images of the Retina; The Vampire Group; Edinburgh United Kingdom) was used to evaluate the retinal vascular calibers according to the VAMPIRE grading protocol. The method is described in details elsewhere (Emanuele Trucco et al., 2015; Perez-Rovira et al., 2011) but will be briefly outlined below.

The software automatically detected retinal vessels and landmarks including the optic disc and the macula, which enabled the creation of a grid around the margin

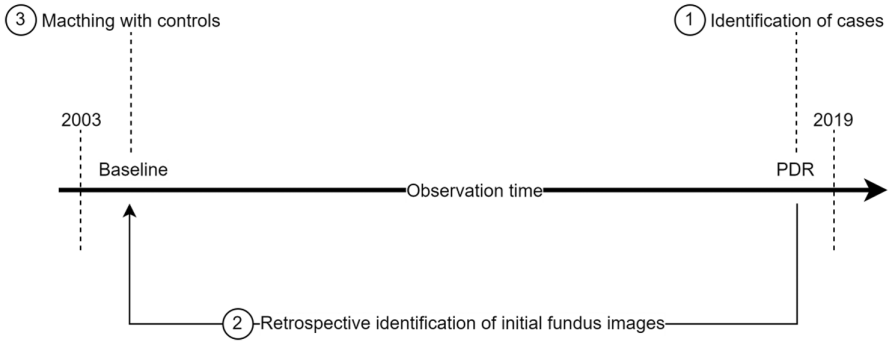


FIGURE 2 A timeline displaying the formation of the study population starting with (1) identification of patients with proliferative diabetic retinopathy in the Funen Diabetes Database from year 2003 to 2019 followed (2) by retrospective identification of initial fundus images and finally (3) matching of cases to controls with the first screening episode as baseline. PDR, proliferative diabetic retinopathy.

TABLE 1 Clinical and demographic differences at the first screening episode for 39 case patients with subsequent progression PDR and 89 control patients with no or minimal (ICDR-level 1) progression in diabetic retinopathy. The patients in both groups had ICDR-level 0 or 1 at baseline.

	Cases	Controls	Total	*p*-Value
Individuals, *n*	39	89	128	
Eyes, *n*	52	107	159	
Baseline DR-level, *n* (%) DR-level 0	33 (63.5%)	63 (59.9%)	96 (60.4%)	
Demographics				
Sex, *n* (%) female	21 (53.8%)	47 (52.2%)	68 (52.7%)	0.87
Age, years (IQR)	42.0 (24.0–60.0)	39.5 (25.0–60.0)	40.0 (25.0–60.0)	0.93
Diabetes type				
1	25 (64.1%)	57 (63.3%)	82 (63.6%)	0.93
2	14 (35.9%)	33 (36.7%)	47 (36.4%)	
Observation time, years (IQR)	10.8 (6.5–11.8)	10.4 (7.2–11.6)	10.5 (7.0–11.7)	0.81
Diabetes duration, years (IQR)	19.0 (12.0–29.0)	14.0 (10.0–21.0)	6.0 (2.0–13.0)	0.01
Para clinical measures				
Body Mass Index, kg/m^2	28.8 (23.9–33.8)	26.7 (24.2–30.1)	27.7 (24.0–31.4)	0.13
Systolic blood pressure, mmHg (IQR)	130 (120–140)	130 (118–137)	130 (120–140)	0.13
Diastolic blood pressure, mmHg (IQR)	80 (74–85)	75 (70–80)	77 (70–82)	0.20
Blood samples				
HbA1c, mmoL/moL (IQR)	73 (60–90)	55 (46–65)	57 (50–75)	<0.001
Serum-creatinine, μmoL/L (IQR)	83.0 (76.0–92.0)	78.5 (66.0–91.0)	80.0 (68.5–91.5)	0.16
LDL, mmoL/L (IQR)	2.45 (2.10–2.90)	2.36 (1.90–2.80)	2.40 (1.90–2.80)	0.73
HDL, mmoL/L (IQR)	1.49 (1.13–2.00)	1.35 (1.08–1.87)	1.39 (1.09–1.88)	0.30
Triglyceride, mmoL/L (IQR)	1.32 (1.03–2.40)	1.16 (0.76–1.69)	1.21 (0.90–1.93)	0.02
Medications				
Insulin, *n* (%) yes	34 (87.2%)	70 (77.8%)	104 (80.6%)	0.21
Retinal geometry				
CRAE, μm (IQR)	229 (211–250)	227 (210–247)	228 (211–248)	0.49
CRVE, μm (IQR)	289 (272–318)	290 (268–312)	290 (270–315)	0.83
AVR	0.78 (0.74–0.85)	0.77 (0.72–0.85)	0.77 (0.72–0.85)	0.86

Note: Continuous variables are presented as medians with interquartile ranges. Categorical variables are presented as numbers with percentages. Case and control differences are calculated at the individual level except for CRAE, CRVE and AVR that are calculated at eye level. Observation time was defined as time from first to last screening or until reach of PDR. Diabetes duration refers to duration of diabetes from debut to last screening episode.

Abbreviations: AVR, arteriole-to-venule ratio; CRAE, central retinal arteriolar equivalent; CRVE, central retinal venular equivalent; DR, diabetic retinopathy; HbA1c, glycosylated haemoglobin; HDL, high-density lipoprotein cholesterol; ICDR, International Clinical Diabetic Retinopathy scale; LDL, low-density lipoprotein cholesterol; PDR, proliferative diabetic retinopathy.

of the optic disc with circular zones A (0.0–0.5 disc diameters from the optic disc margin), B (0.5–1.0) and C (0.5–2.0). The automatic detection was not always on point and manual corrections were made. Retinal vessels were deleted when the vessel type was unclear, two vessels graded as one or if a choroid layer vessel was misinterpreted as a retinal vessel by the software.

The central retinal arteriolar equivalent (CRAE) and central retinal venular equivalent (CRVE) represents a summarization of the six largest arterioles and venules

coursing through zone B. These are automatically calculated in pixels by the VAMPIRE software and has previously been validated by other researchers (Lupascu et al., 2013; McGrory et al., 2018; Project, 2021). We applied conversion from pixels to absolute measurements in microns to ease the interpretation, in which we calculated the conversion factor based on the assumption of an average optic disc diameter of 1800 µm in adult humans (Jonas et al., 1988) and the mean optic disc diameter for all included eyes.

2.6 | Image quality

Every image was manually assessed for quality to exclude ungradable images. Image quality was finally assessed using a binary white and black vessel map as an expression of the software's availability to track the retinal vascularity. Eyes were excluded when the software was unable to track a single quadrant or the majority of the retinal vascularity. No images were excluded due to image quality, which can be related to the fact that patients had no to minimal (ICDR-level 1) when the fundus images were captured.

2.7 | Reliability of retinal vascular analysis

A random sample of 38 (25%) retinal images were included in an intergrader agreement analysis performed between two trained and independent VAMPIRE analysers. The consistency of agreement between the independent analysers was 90% and 94% for CRAE and CRVE, respectively.

2.8 | Statistics

Descriptive statistics are presented as medians with interquartile ranges (continuous data) and numbers with percentages (categorical data). We used the Mann–Whitney *U* test for continuous independent samples and the Chi-square test for categorical independent samples.

We constructed two logistic regression models to test for association between clinical variables (independent) and PDR (dependent): a uni- and a multivariable logistic regression model adjusted for age, sex, HbA1c, blood pressure, duration of diabetes, LDL and triglyceride. The models were performed as mixed-effects regressions with cluster robust standard error to account for the symmetry introduced given that some patients were included with two eyes. We accounted for clustering to address the potential non-independence of observations within each cluster.

Two-way mixed-effects model was used to estimate intraclass correlation coefficients for intergrader agreement on individual measurements for retinal vascular calibers.

Statistical significance was considered with *p*-values under 0.05 and when 1.0 was not included in the 95%-confidence intervals (CI). We performed statistical analyses using STATA17 (StataCorp).

2.9 | Permissions

Approval was given by The Danish Data Agency (journal number: 22/9016) and the study was performed according to the Declaration of Helsinki.

3 | RESULTS

Table 1 shows the baseline characteristics of the 128 individuals included in the study as 39 cases and 89 controls.

Cases had longer duration of diabetes (19 vs. 14 years; $p=0.01$) as well as higher levels of HbA1c (73 vs. 55 mmoL/moL; $p<0.001$) and triglyceride (1.32 vs. 1.16 mmoL/L; $p=0.02$) compared to controls. Cases and controls did not differ according to calibers of retinal arterioles (229 vs. 227 µm; p=0.49), venules (289 vs. 290 µm; $p=0.83$) or the arterio-to-venule ratio (0.78 vs. 0.77; $p=0.86$). Likewise, no differences were found in sex (53.8 vs. 52.2% female; $p=0.87$), age (median 42 vs. 40 years; $p=0.93$), rate of type 1 diabetes (64.1 vs. 63.3%; $p=0.93$), observation time (10.8 vs. 10.4 years; $p=0.81$), body mass index (28.8 vs 26.7 kg/m^2; $p=0.13$), serum-creatinine (83.0 vs. 78.5 µmoL/L; $p=0.16$), LDL (2.45 vs. 2.36 mmoL/L; $p=0.73$), HDL (1.49 vs. 1.35 mmoL/L; $p=0.30$), use of insulin (87.2 vs. 77.8%; $p=0.21$) as well as in systolic (130 vs. 130 mmHg; $p=0.13$) and diastolic blood pressure (80 vs. 75 mmHg; $p=0.20$) between the two groups.

Table 2 shows if independent parameters are associated with PDR in a uni- and a multivariable logistic regression model of 52 case eyes and 107 control eyes. Development of PDR was predicted by HbA1c (OR 1.54 per 10 mmoL/moL; 95%-CI: 1.15–2.07; $p=0.004$), triglyceride (OR 1.39 per 1 mmoL/L; 95%-CI: 1.03–1.86; $p=0.03$) and duration of diabetes (OR 1.09 per 1 year; 95%-CI: 1.03–1.16; $p=0.003$). No other clinical or demographic markers were predictive for upcoming PDR.

4 | DISCUSSION

In this case–control study of patients with no or minimal DR at baseline, long-term development of PDR was predicted by HbA1c, triglycerides, and duration of diabetes, but not by retinal vascular changes.

Previous studies reported that increased CRVE (Klein et al., 2004; Roy et al., 2011; Wong, 2011) predicts incident PDR, whereas CRAE (Cheung et al., 2008; Rogers et al., 2008; Wong, 2011) appears to predict incidence of more mild stages of DR in patients with type 1 diabetes. The findings in the current study did not support these previous findings for CRVE as an early predictor for PDR (64.1% had type 1 diabetes). We were unable to reliably comment on CRAE as a predictor for incident DR due to our selection criteria, in which some patients had mild DR at baseline.

While Klein et al. (Klein et al., 2004) found that larger CRVE was associated with increased 4-, 10- and 14-year PDR incidence in patients with type 1 diabetes, our data suggest otherwise. The difference is likely to be due to the fact that, our cohort consisted of patients with no to mild DR while Klein et al. included

TABLE 2 Uni- and multivariable logistic regression models with odds ratios for association between independent variables and dependent variable PDR. The models include 52 eyes from case patients with subsequent incidence of PDR and 107 eyes from control patients.

Characteristic	Increment	Univariable logistic regression odds ratios (95%-confidence interval)	*p*-Value	Multivariable logistic regression odds ratios (95%-confidence interval)	*p*-Value
Sex	Male vs. female	0.89 (0.40–1.98)	0.78	1.30 (0.45–3.73)	0.63
Age	10 years	0.97 (0.79–1.20)	0.80	0.77 (0.55–1.08)	0.13
Diabetes type	Type 1 vs. 2	1.09 (0.47–2.51)	0.20	1.84 (0.45–7.43)	0.40
Diabetes duration	1 year	1.06 (1.02–1.10)	0.001	1.09 (1.03–1.16)	0.003
Body mass index	1%-point	1.06 (0.99–1.14)	0.07	1.00 (0.89–1.13)	0.94
Systolic blood pressure	5 mmHg	1.09 (0.96–1.23)	0.17	1.18 (0.96–1.45)	0.11
Diastolic blood pressure	5 mmHg	1.20 (0.99–1.46)	0.06	1.00 (0.75–1.33)	0.98
HbA1c	10 mmoL/mol	1.51 (1.22–1.86)	<0.001	1.54 (1.15–2.07)	0.004
Serum-creatinine	1 μmoL/litre	1.01 (0.99–1.03)	0.25	1.02 (0.99–1.04)	0.19
LDL	1 mmoL/litre	1.24 (0.71–2.16)	0.45	1.06 (0.54–2.08)	0.86
HDL	1 mmoL/litre	1.57 (0.82–2.98)	0.17	1.59 (0.65–3.92)	0.31
Triglyceride	1 mmoL/litre	1.69 (1.17–2.44)	0.005	1.39 (1.03–1.86)	0.03
Insulin	Yes vs. no	2.06 (0.67–6.35)	0.21	1.48 (0.32–6.87)	0.62
CRAE	10 μm	0.99 (0.91–1.08)	0.82	1.03 (0.92–1.14)	0.64
CRVE	10 μm	0.99 (0.94–1.05)	0.91	1.01 (0.94–1.10)	0.76
AVR	1 SD	0.95 (0.68–1.33)	0.75	0.98 (0.66–1.46)	0.93

Note: The multivariable model was adjusted for sex, age, HbA1c, diastolic blood pressure, systolic blood pressure, triglyceride and duration of diabetes.
Abbreviations: AVR, arterio-to-venule ratio; CRAE, central retinal arteriolar equivalent; CRVE, central retinal venular equivalent; HbA1c, glycosylated haemoglobin; HDL, high-density lipoprotein cholesterol; LDL, low-density lipoprotein cholesterol; PDR, proliferative diabetic retinopathy; SD, standard deviation.

all patients free of PDR at baseline, meaning that we investigated our population at an earlier stage of DR. The onset of the pathophysiological mechanisms associated with changes in retinal vascular calibre may be related with higher levels of DR. This statement is consistent with CRVE becoming larger as the degree of DR increases reported by several studies (Cheung et al., 2010, 2012; Nguyen et al., 2008).

In a short-term study restricted to African Americans with type-1 diabetes, Roy et al. (Roy et al., 2011) found larger CRVE to be predictive for progression to PDR after 6 years of follow-up. Larger CRVE was found to be a more important PDR predictor in eyes with lower DR grades at baseline, but, again, the patients had higher DR grade than ours (ETDRS ≤53) at baseline (Roy et al., 2011).

In contrast to investigating progression to PDR from all baseline DR levels, Klein et al. and Roy et al. did also investigate the incidence of any DR in patients without DR at baseline, which showed no association with changes in CRAE or CRVE. This is closely comparable to our study suggesting that incident DR or PDR may be difficult to predict with CRVE long before onset of DR in patients with type 1 diabetes (Wong, 2011). In contrast to these findings of CRVE as a predictor for PDR, a long-term study from Denmark in young individuals with type 1 diabetes included with any DR at baseline (Rasmussen et al., 2017), showed that narrower CRAE was associated with 16-year progression to PDR. A pilot study of patients with type 1 and 2 diabetes having minimal DR at baseline (microaneurysms or mild haemorrhages only), found no difference in retinal arteriolar or venular diameter after 4 days of follow-up. This short timeframe is most likely inadequate to develop measurable changes in vessel diameter as a result of vascular endothelial dysfunction (Mlcak et al., 2022).

Only a few studies examined patients with type 2 diabetes exclusively (Cheung et al., 2017; Crosby-Nwaobi et al., 2012). A short-term prospective study by Cheung et al. (Cheung et al., 2017) and a study of 30 eyes by Nwaobi et al. (Crosby-Nwaobi et al., 2012) both reported wider CRAE as predictive for incident PDR while Nwaobi et al. also found that wider CRVE was predictive for PDR in patients with type 2 diabetes who were investigated before onset of any DR.

In comparison to these studies in patients with type 2 diabetes, larger CRAE was associated with progression in DR by Klein et al. (Klein et al., 2004) while narrower CRAE was associated with PDR reported by Broe et al. in type 1 diabetes, which emphasizes that the scientific field maintains a large discrepancy for inexplicable reasons limiting the comparability. These differences include variations in sample size, age, type of diabetes, DR grade at baseline, follow-up periods and potentially in the selection of which semi-automatic computer-assisted software programs was used.

The clinical implementation of CRAE and CRVE as predictors for DR and PDR has so far been unsuccessful due to several challenges. The variety of available semi-automatic software for retinal vascular analysis are time-consuming and, therefore, difficult to implement in a busy clinical setting. Secondly, the lack of clear scientific consensus as well as the small differences of the measured values between study groups makes the clinical interpretation difficult (Crosby-Nwaobi et al., 2012; Klein et al., 2004; Rasmussen et al., 2017; Roy et al., 2011).

We matched the two groups on age, sex, baseline DR-level, type of diabetes and time from first to last screening, which we consider a strength since these are considered risk factors for PDR (Dinesen et al., 2023; Sabanayagam et al., 2019; Wong et al., 2009).

Limitations for the study include that we were unable to find suitable control patients when matching for duration of diabetes leaving the case group with longer duration of diabetes.

In order to be able to match cases and controls in regards to comparability of disease duration and thus a more equal risk-time to develop PDR, we decided to match on time in the screening program as a surrogate measure for duration of diabetes, which is a well-known risk factor for PDR (Song et al., 2018; Yau et al., 2012). We were encouraged to be able to find suitable control patients with this matching criteria and was done to minimize the risk of selecting incomparable controls, which is considered a major weakness in case–control studies.

Examination for signs of diabetic macular oedema (DME) from fundus images is standard of care in the Danish DR screening, but optical coherence tomography is not performed unless it is indicated. Therefore, DME is unfortunately poorly reported in the FDDB and thus we were not able to report the prevalence of DME as a sign of retinal endothelial vascular dysfunction.

We were furthermore unable to calibrate the retinal measurements for axial length, refractive error and corneal curvature as this information is not a part of the Danish DR screening program and is, therefore, absent in the FDDB. The arterio-to-venular ratio is, however, independent of the refraction problem, since the ratio between the two will always be constant regardless of the magnification (Liew et al., 2006).

The calibers of larger vessels in the retina were examined for alterations in relation to diabetes, but it would have been of high interest to study if alterations occur in the even smaller vessels of the retina at this early stage of diabetes. Unfortunately, we were not able to include data from optical coherence tomography angiography or fluorescein angiography due to the register-based nature of the study design that relied on data from the Danish screening program, which only includes six-field fundus imaging as standard examination. Future research should use newer methods, such as optical coherence tomography angiography, to investigate if early capillary dropout increases the risk of long-term progression to PDR.

5 | CONCLUSION

In a long-term study of patients at an early stage of diabetes with no or minimal DR, calibers of the retinal arterioles or venules were not able to predict the incidence of PDR in contrast to well-established risk factors like HbA1c, triglyceride and duration of diabetes.

ACKNOWLEDGEMENTS

The preliminary results of the present study were presented at the 2023 Association for Research in Vision and Ophthalmology (ARVO) in New Orleans, USA.

FUNDING INFORMATION

Steno Diabetes Center Odense, Odense, Denmark and the Region of Southern Denmark, Denmark supported the study. The funding organizations had no role in the design or conduct of this research.

ORCID

Sebastian Dinesen https://orcid.org/0000-0002-7367-9398
Yousif Subhi https://orcid.org/0000-0001-6620-5365
Tunde Peto https://orcid.org/0000-0001-6265-0381
Jakob Grauslund https://orcid.org/0000-0001-5019-0736

REFERENCES

Adelborg, K., Szentkuti, P., Henriksen, J.E., Thomsen, R.W., Pedersen, L., Sundboll, J. et al. (2020) Cohort profile: the Funen diabetes database-a population-based cohort of patients with diabetes in Denmark. *BMJ Open*, 10, e035492.

Carstensen, B., Ronn, P.F. & Jorgensen, M.E. (2020) Prevalence, incidence and mortality of type 1 and type 2 diabetes in Denmark 1996-2016. *BMJ Open Diabetes Research & Care*, 8, e001071.

Cheung, C.Y., Ikram, M.K., Klein, R. & Wong, T.Y. (2015) The clinical implications of recent studies on the structure and function of the retinal microvasculature in diabetes. *Diabetologia*, 58, 871–885.

Cheung, C.Y., Lamoureux, E., Ikram, M.K., Sasongko, M.B., Ding, J., Zheng, Y. et al. (2012) Retinal vascular geometry in Asian persons with diabetes and retinopathy. *Journal of Diabetes Science and Technology*, 6, 595–605.

Cheung, C.Y., Sabanayagam, C., Law, A.K., Kumari, N., Ting, D.S., Tan, G. et al. (2017) Retinal vascular geometry and 6 year incidence and progression of diabetic retinopathy. *Diabetologia*, 60, 1770–1781.

Cheung, N., Mitchell, P. & Wong, T.Y. (2010) Diabetic retinopathy. *Lancet*, 376, 124–136.

Cheung, N., Rogers, S.L., Donaghue, K.C., Jenkins, A.J., Tikellis, G. & Wong, T.Y. (2008) Retinal arteriolar dilation predicts retinopathy in adolescents with type 1 diabetes. *Diabetes Care*, 31, 1842–1846.

Crosby-Nwaobi, R., Heng, L.Z. & Sivaprasad, S. (2012) Retinal vascular calibre, geometry and progression of diabetic retinopathy in type 2 diabetes mellitus. *Ophthalmologica*, 228, 84–92.

Dinesen, S., Stokholm, L., Subhi, Y., Peto, T., Savarimuthu, T.R., Andersen, N. et al. (2023) Five-year incidence of proliferative diabetic retinopathy and associated risk factors in a Nationwide cohort of 201 945 Danish patients with diabetes. *Ophthalmology Science*, 3, 100291.

Emanuele Trucco, A.G., Ballerini, L., Relan, D., Cavinato, A. & MacGillivray, T. (2015) *Morphometric measurements of the retinal vasculature in fundus images with VAMPIRE biomedical image understanding: methods and applications*. Hoboken, NJ: John Wiley & Sons, pp. 91–111.

Ikram, M.K., Cheung, C.Y., Lorenzi, M., Klein, R., Jones, T.L., Wong, T.Y. et al. (2013) Retinal vascular caliber as a biomarker for diabetes microvascular complications. *Diabetes Care*, 36, 750–759.

Ikram, M.K., de Jong, F.J., Bos, M.J., Vingerling, J.R., Hofman, A., Koudstaal, P.J. et al. (2006) Retinal vessel diameters and risk of stroke: The Rotterdam study. *Neurology*, 66, 1339–1343.

Ikram, M.K., de Jong, F.J., van Dijk, E.J., Prins, N.D., Hofman, A., Breteler, M.M. et al. (2006) Retinal vessel diameters and cerebral small vessel disease: The Rotterdam scan study. *Brain*, 129, 182–188.

Ikram, M.K., Witteman, J.C., Vingerling, J.R., Breteler, M.M., Hofman, A. & de Jong, P.T. (2006) Retinal vessel diameters and risk of hypertension: The Rotterdam study. *Hypertension*, 47, 189–194.

Jonas, J.B., Gusek, G.C. & Naumann, G.O. (1988) Optic disc, cup and neuroretinal rim size, configuration and correlations in normal eyes. *Investigative Ophthalmology & Visual Science*, 29, 1151–1158.

Klein, R., Klein, B.E., Moss, S.E., Wong, T.Y., Hubbard, L., Cruickshanks, K.J. et al. (2004) The relation of retinal vessel caliber to the incidence and progression of diabetic retinopathy: XIX: The Wisconsin epidemiologic study of diabetic retinopathy. *Archives of Ophthalmology*, 122, 76–83.

Larsen, M.B., Henriksen, J.E., Grauslund, J. & Peto, T. (2017) Prevalence and risk factors for diabetic retinopathy in 17 152 patients from the Island of Funen, Denmark. *Acta Ophthalmologica*, 95, 778–786.

Liew, G., Mitchell, P., Wang, J.J. & Wong, T.Y. (2006) Effect of axial length on retinal vascular network geometry. *American Journal of Ophthalmology*, 141, 597–598.

Lupascu, C.A., Tegolo, D. & Trucco, E. (2013) Accurate estimation of retinal vessel width using bagged decision trees and an extended multiresolution Hermite model. *Medical Image Analysis*, 17, 1164–1180.

McGrory, S., Taylor, A.M., Pellegrini, E., Ballerini, L., Kirin, M., Doubal, F.N. et al. (2018) Towards standardization of quantitative retinal vascular parameters: comparison of SIVA and VAMPIRE measurements in the Lothian birth cohort 1936. *Translational Vision Science & Technology*, 7, 12.

Mlcak, P., Chlup, R., Kudlova, P., Krystynik, O., Kral, M., Kucerova, V. et al. (2022) Retinal oxygen saturation is associated with HbA1c but not with short-term diabetes control, internal environment, smoking and mild retinopathy - ROXINEGLYD study. *Acta Ophthalmologica*, 100, e142–e149.

Nguyen, T.T., Wang, J.J., Sharrett, A.R., Islam, F.M., Klein, R., Klein, B.E. et al. (2008) Relationship of retinal vascular caliber with diabetes and retinopathy: the multi-ethnic study of atherosclerosis (MESA). *Diabetes Care*, 31, 544–549.

Perez-Rovira, A., MacGillivray, T., Trucco, E., Chin, K.S., Zutis, K., Lupascu, C. et al. (2011) VAMPIRE: vessel assessment and measurement platform for images of the REtina. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, 3391–3394.

Project I. (2021) On the quantitative effects of compression of retinal fundus images on morphometric vascular measurements in VAMPIRE. *Computer Methods and Programs in Biomedicine*, 202, 105969.

Rasmussen, M.L., Broe, R., Frydkjaer-Olsen, U., Olsen, B.S., Mortensen, H.B., Peto, T. et al. (2017) Retinal vascular geometry and its association to microvascular complications in patients with type 1 diabetes: the Danish cohort of pediatric diabetes 1987 (DCPD1987). *Graefe's Archive for Clinical and Experimental Ophthalmology*, 255, 293–299.

Rogers, S.L., Tikellis, G., Cheung, N., Tapp, R., Shaw, J., Zimmet, P.Z. et al. (2008) Retinal arteriolar caliber predicts incident retinopathy: the Australian diabetes, obesity and lifestyle (AusDiab) study. *Diabetes Care*, 31, 761–763.

Roy, M.S., Klein, R. & Janal, M.N. (2011) Retinal venular diameter as an early indicator of progression to proliferative diabetic retinopathy with and without high-risk characteristics in African Americans with type 1 diabetes mellitus. *Archives of Ophthalmology*, 129, 8–15.

Sabanayagam, C., Banu, R., Chee, M.L., Lee, R., Wang, Y.X., Tan, G. et al. (2019) Incidence and progression of diabetic retinopathy: a systematic review. *The Lancet Diabetes and Endocrinology*, 7, 140–149.

Song, P., Yu, J., Chan, K.Y., Theodoratou, E. & Rudan, I. (2018) Prevalence, risk factors and burden of diabetic retinopathy in China: a systematic review and meta-analysis. *Journal of Global Health*, 8, 010803.

Wilkinson, C.P., Ferris, F.L., 3rd, Klein, R.E., Lee, P.P., Agardh, C.D., Davis, M. et al. (2003) Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*, 110, 1677–1682.

Wong, T.Y. (2011) Retinal vessel diameter as a clinical predictor of diabetic retinopathy progression: time to take out the measuring tape. *Archives of Ophthalmology*, 129, 95–96.

Wong, T.Y., Klein, R., Couper, D.J., Cooper, L.S., Shahar, E., Hubbard, L.D. et al. (2001) Retinal microvascular abnormalities and incident stroke: the atherosclerosis risk in communities study. *Lancet*, 358, 1134–1140.

Wong, T.Y., Mwamburi, M., Klein, R., Larsen, M., Flynn, H., Hernandez-Medina, M. et al. (2009) Rates of progression in diabetic retinopathy during different time periods: a systematic review and meta-analysis. *Diabetes Care*, 32, 2307–2313.

Yau, J.W., Rogers, S.L., Kawasaki, R., Lamoureux, E.L., Kowalski, J.W., Bek, T. et al. (2012) Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*, 35, 556–564.

How to cite this article: Dinesen, S., Stokholm, L., Subhi, Y., Henriksen, J.E., Savarimuthu, T.R., Peto, T. et al. (2024) Retinal main vessel calibers and systemic markers for long-term development of proliferative diabetic retinopathy. *Acta Ophthalmologica*, 102, 448–454. Available from: https://doi.org/10.1111/aos.15780

Prediction of Proliferative Diabetic Retinopathy from Retinal Images using Deep Learning

Sebastian Dinesen, MD[1, 2, 3], Lonny Stokholm, PhD[4], Andrea Blasi Núñez, cand. polyt.[5], Jakob Kristian Holm Andersen, PhD[5], Kristian Nielsen, cand. polyt.[6], Jacob Volmer Stidsen, PhD[3], Thiusius Rajeeth Savarimuthu, PhD[5], Tunde Peto, DMSc[7], Jakob Grauslund, DMSc[1, 2, 3]

Affiliations

[1]Department of Ophthalmology, Odense University Hospital, Odense, Denmark,

[2]Department of Clinical Research, University of Southern Denmark, Odense, Denmark,

[3]Steno Diabetes Center Odense, Odense University Hospital, Odense, Denmark,

[4]Open Patient Data Explorative Network, Department of Clinical Research, University of Southern Denmark and Odense University Hospital, Odense, Denmark

[5]The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark.

[6]Department of Political Science and Public Management, University of Southern Denmark, Odense, Denmark.

[7]Centre for Public Health, Queen's University Belfast, Belfast, UK

Address for correspondence

Dr. Sebastian Dinesen
Department of Ophthalmology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense, Denmark.
E-mail: sebastian.dinesen@rsyd.dk | Phone: +45 2623 7090 | Fax: +45 6541 3004

Abstract

Objective

We aimed to develop and validate a deep learning algorithm to predict future development of proliferative diabetic retinopathy (PDR) from six-field retinal images. The study also aimed to develop prediction models from baseline DR levels 1 to 3.

Design

A diagnostic accuracy study using deep learning on retrospectively collected retinal images.

Subjects

We included retinal images from patients with type 1 or type 2 diabetes attending the Danish screening program for diabetic retinopathy (DR).

Methods

We included retinal images from the Danish Registry of Diabetic Retinopathy (DiaBase), which is a Danish national DR screening program containing the DR level, date of screening and more from each clinical screening episode. The patients registered with PDR in the DiaBase were included as case patients while those with DR level 0 to 3 as control patients. All images prior to the diagnosis of PDR were included for the cases while all available images for the controls were included. We split the images into training (70%), validation (10%) and testing (20%) datasets. The development of the algorithm included a multi-step approach, which included vessel segmentation with the FR-Unet model to enhance feature representation, numerical vessel features density and tortuosity and extraction of features from raw images.

Main Outcome Measures

Sensitivity, specificity, positive predictive value and negative predictive value for prediction of PDR.

Results

We included 26,828 retinal images in the first model that comprised all DR levels with 627 case images and 26,201 control images. The deep learning model with prediction from ICDR level 1 included 3,672 images with 22 cases and 3,650 controls. The model based on ICDR level 2 included 5,867 images with 208 cases and 5,659 controls. The model based on ICDR level 3 included 1,693 images with 376 cases and 1,317 controls. None of the models succeeded to perform any positive predictive values equal to 0% for all main outcome measures.

Conclusion

This study did not succeed in the prediction of PDR from a dataset of 26,828 retinal images. It outlines methodological challenges and potential solutions related to class imbalance, limited data, and longitudinal data with fixed time points.

Introduction

The global burden of diabetes continues to rise with an estimation of 828 million adults affected in 2022, which is an increase of 630 million since 1990[1]. Diabetic retinopathy (DR) is a well-known microvascular complication to diabetes, with its sight-threatening end-stage, proliferative diabetic retinopathy (PDR), affecting 32.4% of the patients with type 1 diabetes and 3.0% of those with type 2 diabetes in 2012[2]. While the incidence of PDR has shown a downward trend in Western countries[3], the overall demand for DR screening is expected to rise due to the increasing prevalence of diabetes. To manage this growing workload, it is crucial to allocate medical resources toward individuals at risk of progressing to PDR. This approach is supported by studies that generally advocate for extended screening intervals for a large proportion of low-risk patients, with minimal risk of delayed PDR detection[4-6]. Despite these recommendations, it has been challenging to extend screening intervals due to difficulties in predicting an individual's risk of and time to progression. A major limitation of current DR screening programs is the lack of individualized risk prediction and in particular the identification of patients at high risk of developing sight-threatening DR. The risk of progression is influenced by both non-modifiable and modifiable factors and varies considerably between individuals[7]. Current screening strategies rely primarily on population-based intervals rather than personalized assessments, which may result in over-screening of low-risk individuals.

Artificial intelligence (AI) has shown great potential and promising results in addressing challenges within healthcare systems worldwide[8]. Deep learning (DL), a branch of AI, utilizes artificial neural networks that mimic the human central nervous system and holds great potential for healthcare solutions. Convolutional neural networks (CNN) are a subclass of DL models designed to detect patterns and features in visual data[9]. Current DL models in the field of DR primarily focus on detecting existing disease[10-15], while only a few aim to predict the onset or progression of DR[6,12].

We hypothesized that progression to PDR could be predicted from retinal images obtained before the date of PDR diagnosis. The purpose of this study was to develop a DL model to predict progression to PDR from all DR levels combined and to create separate predictive DL models for retinal images stratified by DR levels 1 to 3.

Methods

<u>Design and Setting</u>

This is a DL diagnostic accuracy study based on retrospective register data. We included patients from the Southern Region of Denmark screened at a public hospital facility from January 2013 to December 2022.

Study Population

The study population included individuals with type 1 or type 2 diabetes identified through the Danish Registry of Diabetic Retinopathy (DiaBase)[16]. The patients with a PDR diagnosis in the DiaBase were included as case patients while those with levels of DR from 0 to 3 were included as control patients.

Danish Health Registries

The Civil Registration System (CRS)

Since April 2, 1968, every child born to a registered mother in Denmark has been assigned a ten-digit Civil Personal Register (CPR) number. This identifier is used across all Danish medical and administrative databases, which makes it essential for linking information across registers. The Central Population Register (CRS) includes variables such as date of birth, sex, citizenship, civil status, migration history, and vital status.

The Danish Registry of Diabetic Retinopathy (DiaBase)

The DiaBase is a nationwide clinical quality database that collects data from individuals with type 1 and type 2 diabetes who participate in the Danish DR screening program[16]. The database records the severity of DR as graded by ophthalmologists based on the International Clinical Diabetic Retinopathy Severity Scale (ICDR), ranging from level 0 (no DR) to level 4 (proliferative DR). The database also includes screening dates, recommended follow-up intervals, visual acuity, and screening indications.

Identification of Retinal Images

Retinal images from the DR screening program in the Southern Region of Denmark, tagged with the CPR number and screening date, are stored in a local database at Odense University Hospital (OUH). We included all available images captured prior to the date of PDR diagnosis for the case group, regardless of how long before the diagnosis they were obtained, and excluded any images captured after the diagnosis. The retinal images from those without a PDR diagnosis were included as controls. We included all levels of DR for model development, while also creating smaller datasets stratified by DR levels 1, 2, and 3 to develop individual PDR prediction DL models for each (Table 1). The images were allocated for training (70%), validation (10%) and testing (20%) datasets. Each individual could contribute several images, but only represent one of the three datasets.

Algorithm Development

We designed a multi-step pipeline to enhance the model's ability to predict progression to PDR by incorporating both image-based and vessel-specific features. The first stage involved segmenting the retinal vasculature from each image using the FR-UNet model[17], a DL architecture specialized for segmenting biological structures. Trained on datasets containing annotated vessel images, FR-UNet generated vessel masks highlighting the vascular network. These masks were then integrated as an additional input channel alongside the original RGB image to enrich the input data with explicit vascular information.

We utilized the InceptionV3 CNN[18] for feature extraction, which originally was pre-trained on the ImageNet dataset for large-scale image classification tasks. InceptionV3 is capable of capturing both fine details and broader spatial structures within images. To tailor the model to our classification task, we adopted a transfer learning approach where the lower layers, responsible for general feature extraction, were kept frozen, while only the final layers were fine-tuned using our dataset.

We calculated numerical vessel-related metrics to further guide the model in addition to deep features extracted from the retinal images and vessel mask. These included vessel density that was computed as the proportion of vessel pixels relative to the total image area

$$\text{Vessel density} = \frac{\text{Number of vessel pixels}}{\text{Total image pixels}}$$

We also assessed vessel tortuosity by first generating a skeletonized representation of the segmented vessels (Figure 2). By using this one-pixel-wide vessel map, we identified bifurcation points and defined vessel segments as paths between consecutive bifurcations, vessel endpoints, or up to a predefined maximum length. For each segment, tortuosity was calculated as the ratio of the actual path length to the straight-line distance between endpoints:

$$\text{Tortuosity} = \frac{\text{Segment length}}{\text{Euclidean Distance}}$$

We now extracted the maximum, minimum, and mean tortuosity values across all segments in order to summarize the tortuosity at the image level. These numerical features were processed separately through a fully connected multi-layer perceptron (MLP).

The deep features from the InceptionV3 network and numerical features from the MLP were combined into a single representation. This joint feature vector was passed through a final fully connected classification layer to generate the prediction. Model training was performed using binary cross-entropy loss that we optimized with the Adam optimizer[19]. Due to class imbalance in our image dataset, with fewer samples representing the PDR group, we applied class weighting to the loss function to increase the model's sensitivity toward the minority class[20].

Results

We included 26,828 retinal images in the first model comprising all DR levels divided as 627 case images and 26,201 control images. The DL model with prediction from ICDR level 1 included 3,672 images allocated as 22 case and 3,650 control images, while we included 5,867 images to the DL model with images stratified to ICDR level 2 allocated as 208 case and 5,659 control images. The last DL model predicting from ICDR level 3 included 376 case images and 1,317 control images equal to 1,693 images in total (Table 1).

All of the Dl models failed to perform any positive predictions resulting in sensitivities, specificities, positive predictive values and negative predictive values of 0% (Table 2).

Discussion

This study utilized 26,828 images to develop and test a DL model for predicting future progression to PDR. The model failed to generate any positive predictions, likely due to the limited number of case images and the inherent challenges of a relatively small dataset. As a result, the extracted image and vascular geometry features may have contained insufficient predictive signal.

Several studies have explored the potential of DL models to predict systemic conditions and patient characteristics from retinal images, but many of these were cross-sectional predictions of factors such as HbA1c levels, kidney disease or sex[21,22]. While these findings demonstrated the rich information embedded in retinal images, their clinical utility remains limited, as many of these conditions are easily assessed through standard laboratory tests. Fewer studies have attempted to predict future disease progression, which holds greater clinical relevance by enabling early interventions and personalized care. Dai et al published a paper in 2024 presenting a study in which they succeeded to predict individualized time to DR progression over a five-year period using 717,308 retinal images for pre-training and 118,868 baseline retinal images for model

development[23]. They predicted the one to five year risk of vision-threatening DR (defined as ICDR level 3 to ICDR level 4) with AUCs ranging from 0.825 to 0.860 from retinal images alone. They also predicted the time to progression of DR, which extended screening intervals from 12 to 32 months, resulting in a 62.5% reduction in screening frequency with only a 0.18% delayed VTDR detection.

Another study[24] developed and tested two versions of a DL system (one field image model and a three field image model) to predict the onset of DR within two years in patients without DR at baseline. The models were trained on 575,431 images and evaluated through internal (3,678 images) and external (2,345 images) testing. The AUC was 0.77 for the one-field model and 0.79 for the three-field model. Adding baseline characteristics such as HbA1c, duration of diabetes, self-reported diabetic control, and insulin use did not significantly improve performance. The clinical value of the system is limited, as patients with diabetes are expected to develop DR at some point over time, and the model does not support extending screening intervals beyond two years, which was the study's follow-up period.

Developing DL models for rare outcomes like PDR is challenging because CNNs require large amounts of data to learn disease-specific features. Although methods exist to address small data challenges, insufficient data remains a reasonable explanation for the limited performance of our proposed model[25]. We included 26,828 images, whereas the study by Dai et al[23] used a total nearly one million images for pre-training, development and validation of their DeepDR Plus model. Their approach to predict vision-threatening DR is comparable to ours, although they defined vision-threatening DR as ICDR levels 3 and 4, while we focused specifically on predicting PDR (ICDR level 4). They used baseline retinal images to predict vision-threatening DR and also developed a combined model incorporating clinical and demographic data. Their findings showed no improvement in performance when adding these additional data, which indicates that the predictive information is already contained within the retinal images. The few PDR cases compared to controls, which introduces class imbalance may have been a major challenge. Dai et al. did on the other hand not include overwhelming number of PDR cases either, as they reported 681 eyes with progression to PDR and 1,382 eyes with progression to severe non-PDR compared to 627 PDR images in our study. Therefore, pre-training on a large dataset of retinal images may be valuable for optimizing feature extraction, even though both our study (using InceptionV3) and the study by Dai et al. (using ResNet-50) employed transfer learning from a pre-trained CNN as a backbone.

The severe class imbalance was a major challenge in this study, as the primary outcome was to predict PDR and only a small fraction of images represented patients who progressed. Weighting the loss function is a common method to address class imbalance by assigning higher penalties to errors on the minority class as

this will force the model to focus more on rare cases during training. However, the imbalance in this study may have been so severe that even weighted loss could not compensate for the limited number of PDR cases. With so few progression events and an overall limited data volume, the model may have struggled to learn meaningful patterns associated with future PDR.

Our study was limited by the cross-sectional nature of the data and the absence of fixed follow-up time points. Ideally, the dataset would include a baseline image from each eye and track progression at predefined intervals, such as 3 months to 5 years. Incorporating cases with long-term follow-up is important in a chronic disease like DR because the progression often occurs slowly over time.

Future studies should shift focus toward the prediction of less rare outcomes, but with equal clinical relevance as that introduced with the PDR prediction, such as time to DR progression, which offers a more individualized approach and allows every patient in longitudinal datasets to contribute to the outcome. This strategy may help overcome challenges related to small data and severe class imbalance. To further overcome small data issues, incorporation of baseline clinical and demographic data, such as HbA1c levels, age, duration of diabetes, type of diabetes and more, could potentially strengthen the predictive performance even though the DL systems in the studies by Dai et al[23] and Bora et al[24] did not improve much by adding complementary data.

Conclusion

This study did not succeed in predicting PDR from a dataset of 26,828 retinal images. It highlights important methodological challenges and presents potential solutions, such as handling class imbalance, small data problems and managing longitudinal data structures with fixed time points for future research to build upon.

References

1. Zhou B, Rayner AW, Gregg EW, et al. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. *The Lancet*. 2024;404(10467):2077-2093. doi:10.1016/S0140-6736(24)02317-1
2. Yau JW, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*. Mar 2012;35(3):556-64. doi:10.2337/dc11-1909
3. Abou Taha A, Dinesen S, Vergmann AS, Grauslund J. Present and future screening programs for diabetic retinopathy: a narrative review. *Int J Retina Vitreous*. Feb 3 2024;10(1):14. doi:10.1186/s40942-024-00534-8
4. Dinesen S, Stokholm L, Subhi Y, et al. Five-Year Incidence of Proliferative Diabetic Retinopathy and Associated Risk Factors in a Nationwide Cohort of 201 945 Danish Patients with Diabetes. *Ophthalmol Sci*. Sep 2023;3(3):100291. doi:10.1016/j.xops.2023.100291
5. Grauslund J, Stokholm L, Andersen N, et al. Risk of 2-year progression from no diabetic retinopathy to proliferative diabetic retinopathy in accordance with glycaemic regulation in 145 527 persons with diabetes in a national cohort. *Acta Ophthalmol*. May 2024;102(3):e407-e409. doi:10.1111/aos.15806
6. Zhang K, Liu X, Xu J, et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat Biomed Eng*. Jun 2021;5(6):533-545. doi:10.1038/s41551-021-00745-6
7. Jenkins AJ, Joglekar MV, Hardikar AA, Keech AC, O'Neal DN, Januszewski AS. Biomarkers in Diabetic Retinopathy. *Rev Diabet Stud*. Spring-Summer 2015;12(1-2):159-95. doi:10.1900/rds.2015.12.159
8. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. *Brief Bioinform*. Nov 27 2018;19(6):1236-1246. doi:10.1093/bib/bbx044
9. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. May 28 2015;521(7553):436-44. doi:10.1038/nature14539
10. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med*. 2018;1:39. doi:10.1038/s41746-018-0040-6
11. Chia MA, Hersch F, Sayres R, et al. Validation of a deep learning system for the detection of diabetic retinopathy in Indigenous Australians. *Br J Ophthalmol*. Jan 29 2024;108(2):268-273. doi:10.1136/bjo-2022-322237
12. Dai L, Wu L, Li H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nat Commun*. May 28 2021;12(1):3242. doi:10.1038/s41467-021-23458-5
13. Ipp E, Liljenquist D, Bode B, et al. Pivotal Evaluation of an Artificial Intelligence System for Autonomous Detection of Referrable and Vision-Threatening Diabetic Retinopathy. *JAMA Netw Open*. Nov 1 2021;4(11):e2134254. doi:10.1001/jamanetworkopen.2021.34254
14. Li Z, Keel S, Liu C, et al. An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs. *Diabetes Care*. Dec 2018;41(12):2509-2516. doi:10.2337/dc18-0147
15. Ting DSW, Cheung CY, Lim G, et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *Jama*. Dec 12 2017;318(22):2211-2223. doi:10.1001/jama.2017.18152
16. Andersen N, Hjortdal J, Schielke KC, et al. The Danish Registry of Diabetic Retinopathy. *Clin Epidemiol*. 2016;8:613-619. doi:10.2147/clep.S99507
17. Liu W, Yang H, Tian T, et al. Full-Resolution Network and Dual-Threshold Iteration for Retinal Vessel and Coronary Angiograph Segmentation. *IEEE J Biomed Health Inform*. Sep 2022;26(9):4623-4634. doi:10.1109/jbhi.2022.3188710

18. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016; https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.308
19. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *CoRR*. 2014;abs/1412.6980
20. Buda M, Maki A, Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw*. 2018;106(C):249–259. doi:10.1016/j.neunet.2018.07.011
21. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. Mar 2018;2(3):158-164. doi:10.1038/s41551-018-0195-0
22. Sabanayagam C, Xu D, Ting DSW, et al. A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. *Lancet Digit Health*. Jun 2020;2(6):e295-e302. doi:10.1016/s2589-7500(20)30063-7
23. Dai L, Sheng B, Chen T, et al. A deep learning system for predicting time to progression of diabetic retinopathy. *Nat Med*. Feb 2024;30(2):584-594. doi:10.1038/s41591-023-02702-z
24. Bora A, Balasubramanian S, Babenko B, et al. Predicting the risk of developing diabetic retinopathy using deep learning. *The Lancet Digital Health*. 2021;3(1):e10-e19. doi:10.1016/S2589-7500(20)30250-8
25. Rather IH, Kumar S, Gandomi AH. Breaking the data barrier: a review of deep learning techniques for democratizing AI with small datasets. *Artificial Intelligence Review*. 2024/08/02 2024;57(9):226. doi:10.1007/s10462-024-10859-3

Figure Legends

Figure 1

The figure illustrates the retrospective identification of early retinal images (1) from patients who later developed proliferative diabetic retinopathy (PDR). The six-field retinal images were collected prior to the PDR event and deep learning was applied to these early images (2) to predict future PDR development (3).

Figure 2

The figure illustrates four examples of vessel segmentation applied to retinal images using the FR-Unet model.

Table 1: The Number of Images Dichotomized into Training, Validation and Testing Datasets for each of the four deep learning prediction models.

	All		ICDR level 1		ICDR level 2		ICDR level 3	
Subset, n	Control	Cases	Control	Cases	Control	Cases	Control	Cases
Training	19,812	462	2,746	17	4,373	146	1,001	284
Validation	2,594	95	352	2	556	39	116	54
Testing	3,795	70	552	3	730	23	200	38
Total, n	26,201	627	3,650	22	5,659	208	1,317	376

ICDR = International Clinical Diabetic Retinopathy Severity Scale.

Table 2: Performance of the Four Deep Learning Models for Predicting Future PDR

%	All	ICDR level 1	ICDR level 2	ICDR level 3
Sensitivity	0	0	0	0
Specificity	0	0	0	0
PPV	0	0	0	0
NPV	0	0	0	0

PDR = proliferative diabetic retinopathy. ICDR = International Clinical Diabetic Retinopathy Severity Scale. All refers to all diabetic retinopathy levels combined.

Figure 1:

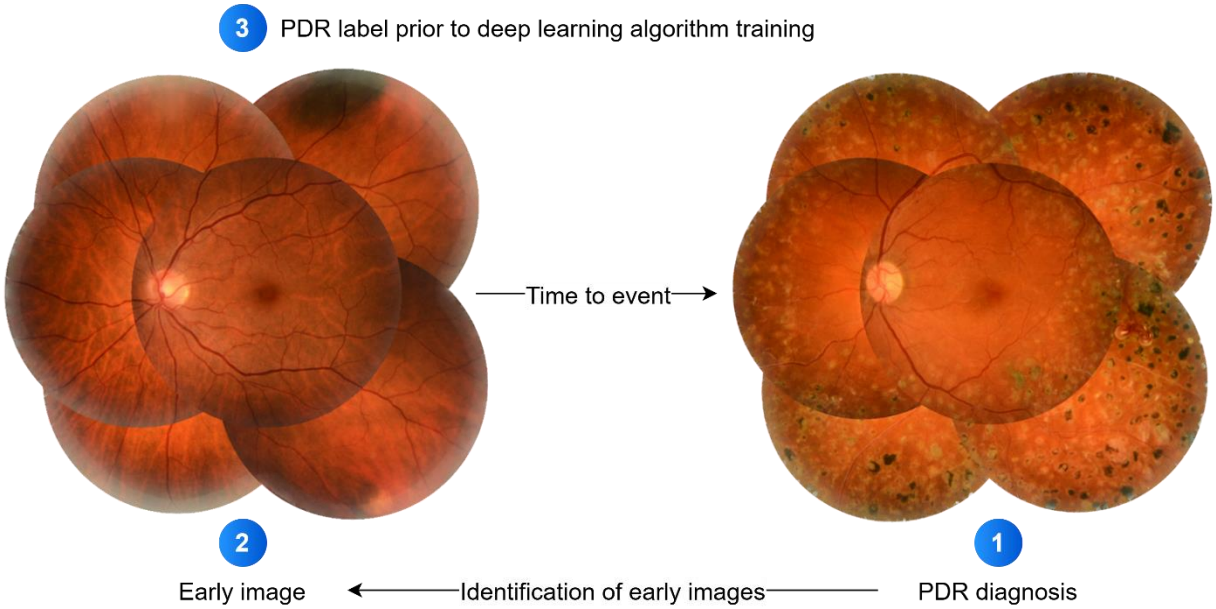
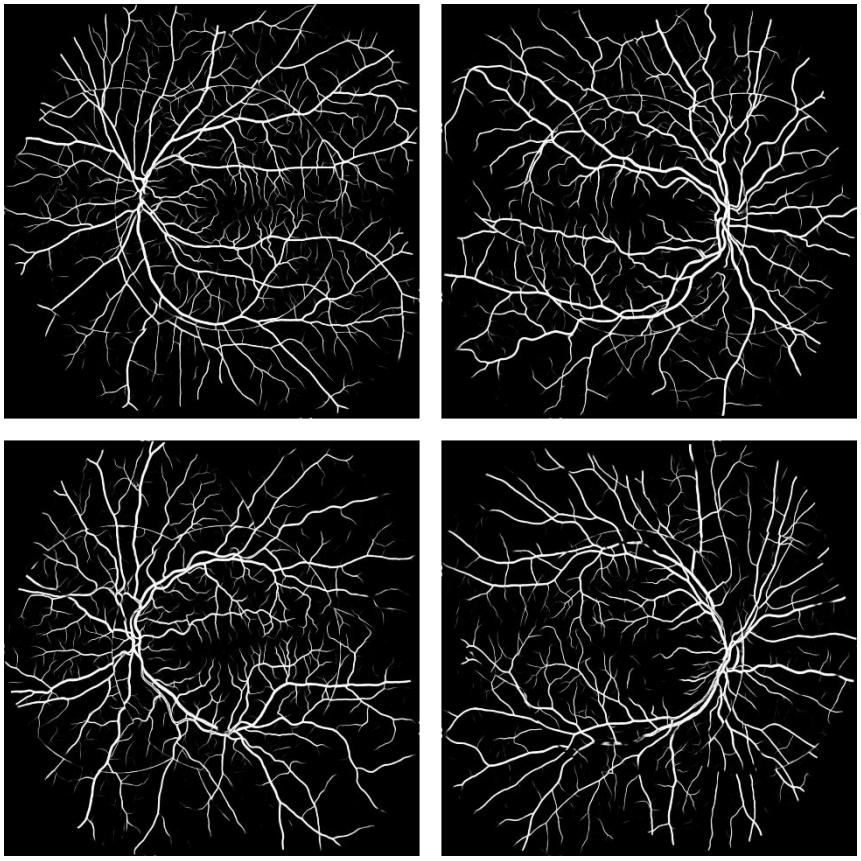

3
PDR label prior to deep learning algorithm training
Time to event
2
1
Early image
Identification of early images
PDR diagnosis

Figure 2:



Ophthalmol Ther
https://doi.org/10.1007/s40123-025-01127-w

ORIGINAL RESEARCH

A Deep Learning Segmentation Model for Detection of Active Proliferative Diabetic Retinopathy

Sebastian Dinesen · Marianne G. Schou · Christoffer V. Hedegaard · Yousif Subhi · Thiusius R. Savarimuthu · Tunde Peto · Jakob K. H. Andersen · Jakob Grauslund

Received: February 14, 2025 / Accepted: March 5, 2025

ABSTRACT

***Introduction*:** Existing deep learning (DL) algorithms lack the capability to accurately identify patients in immediate need of treatment for proliferative diabetic retinopathy (PDR). We aimed to develop a DL segmentation model to detect active PDR in six-field retinal images by the annotation of new retinal vessels and preretinal hemorrhages.

***Methods*:** We identified six-field retinal images classified at level 4 of the International Clinical Diabetic Retinopathy Disease Severity Scale collected at the Island of Funen from 2009 to 2019 as part of the Danish screening program for diabetic retinopathy (DR). A certified grader (grader 1) manually dichotomized the images into active or inactive PDR, and the images were then reassessed by two independent certified graders. In cases of disagreement, the final classification decision was made in collaboration between grader 1 and one of the secondary graders. Overall, 637 images were classified as active PDR. We then applied our pre-established DL segmentation model to annotate nine lesion types before training the algorithm. The segmentations of new vessels and preretinal hemorrhages were corrected for any inaccuracies before training the DL algorithm. After the classification and pre-segmentation phases the images were divided into training (70%), validation (10%), and testing (20%) datasets. We added 301 images with inactive PDR to the testing dataset.

***Results*:** We included 637 images of active PDR and 301 images of inactive PDR from 199 individuals. The training dataset had 1381 new vessel and preretinal hemorrhage lesions, while the validation dataset had 123 lesions and the testing dataset 374 lesions. The DL system demonstrated a sensitivity of 90% and a specificity

S. Dinesen (✉) · M. G. Schou · C. V. Hedegaard · J. Grauslund
Department of Ophthalmology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense, Denmark
e-mail: sebastian.dinesen@rsyd.dk

S. Dinesen · M. G. Schou · Y. Subhi · J. Grauslund
Department of Clinical Research, University of Southern Denmark, Odense, Denmark

S. Dinesen · J. Grauslund
Steno Diabetes Centre Odense, Odense University Hospital, Odense, Denmark

Y. Subhi
Department of Ophthalmology, Rigshospitalet, Glostrup, Denmark

Y. Subhi
Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

T. R. Savarimuthu · J. K. H. Andersen
The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

T. Peto
Centre for Public Health, Queen's University Belfast, Belfast, UK

Published online: 27 March 2025

of 70% for annotation-assisted classification of active PDR. The negative predictive value was 94%, while the positive predictive value was 57%.

Conclusions: Our DL segmentation model achieved excellent sensitivity and acceptable specificity in distinguishing active from inactive PDR.

Keywords: Deep learning; Proliferative diabetic retinopathy; Black box elimination; New vessels; Preretinal hemorrhages

Key Summary Points

Why carry out this study?
A deep learning (DL) model capable of identifying patients in urgent need of treatment and distinguishing active proliferative diabetic retinopathy (PDR) from stable inactive PDR could help optimize resource allocation and improve patient care.
Existing DL models lack the ability to accurately differentiate between active and inactive PDR.
Can a DL segmentation model accurately identify active PDR by detecting new retinal vessels and preretinal hemorrhages?
What was learned from the study?
The DL segmentation model demonstrated high sensitivity in detecting active PDR, which makes it a promising tool for screening and clinical decision support.
The high sensitivity and negative predictive value suggest that the model could be valuable in ruling out inactive cases and potentially improve screening workflows.
Further optimization is needed to enhance specificity and positive predictive value to ensure more precise identification of patients requiring urgent treatment.

INTRODUCTION

The current prevalence of patients with diabetes is 463 million and increases worldwide [1] and the associated systemic complications, including diabetic retinopathy (DR), are projected to remain high [2]. Proliferative diabetic retinopathy (PDR) develops from chronic hyperglycemia, which causes endothelial dysfunction and capillary occlusion, leading to retinal ischemia. This triggers the release of the proangiogenic vascular endothelial growth factor (VEGF) to form new vulnerable and potentially sight-threatening vessels in the retina [3].

The global prevalence of PDR is estimated to be 32.4% in patients with type 1 and 3.0% in type 2 diabetes [4]. Screening for DR and improved management of systemic factors has decreased the risk of vision loss from PDR [5, 6], but the rising prevalence of diabetes increases the number of patients at risk of DR, which underscores the need for continued screening efforts.

Deep learning (DL) is a machine learning technique particularly suited for image recognition that has shown promising diagnostic potential [7–9]. While classification models assign a diagnosis without revealing their reasoning, DL segmentation models go further by highlighting pathological lesions directly within the image [10]. This added transparency helps clinicians see the model's decision-making process and better interpret the results (Fig. 1).

Several research groups have developed DL models demonstrating excellent performance in detecting sight-threatening DR with sensitivity and specificity rates exceeding 90% [11–17], while human graders achieve a sensitivity of 88.5% and a specificity of 99.6% [11]. However, these DL models are trained to identify those with International Clinical Diabetic Retinopathy Disease Severity Scale (ICDR) level 2 or more as screening positive. This motivated us to develop a DL algorithm to detect eyes with active PDR that require immediate treatment, as moderate NPDR does not need treatment intervention. In many low- to middle-income countries with limited healthcare resources, referring mild NPDR cases would overburden

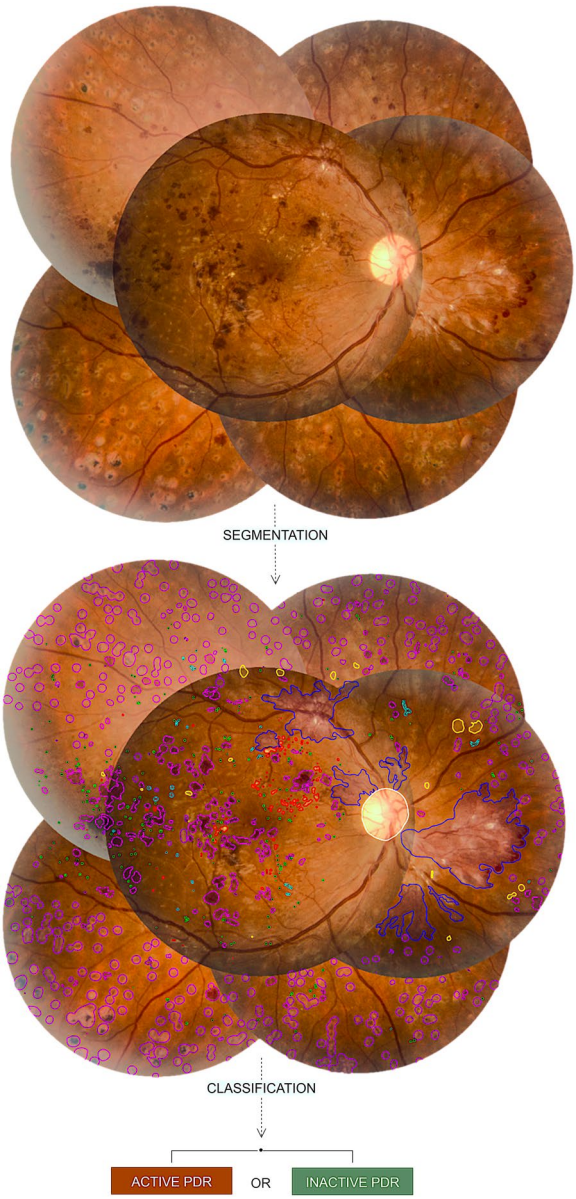


Fig. 1 The figure illustrates the objective of the study presenting an example of an original six-field retinal image and the same image with pre-segmentation of new vessels (dark blue), preretinal hemorrhages (not present), peripheral laser tracks (purple), central laser tracks (light blue), intraretinal hemorrhages (pink), microaneurysms (green), soft exudates (yellow), hard exudates (red), and intraretinal microvascular abnormalities (turquoise) before training of the new model to distinguish active from inactive PDR. Certified grader 1 evaluated the segmentations of all new vessel in the presented image and reannotated all of the pre-segmentations made by the DL system to ensure complete precision. *PDR* proliferative diabetic retinopathy, *DL* deep learning

screening systems. A DL system that identifies those in urgent need of treatment could help optimize these resources. Our team has previously developed a DL model that classifies DR on the ICDR scale from 0 to 4 and annotates eight specific DR lesions [10]. This model detected ICDR level 4 with an accuracy of 96.2%, but it did not differentiate between already treated inactive PDR and referable active PDR. It would be of high value for a DL model to separate those with active PDR from those with stable inactive PDR to ensure that healthcare resources are directed toward patients requiring urgent intervention.

We aimed to develop and validate a DL segmentation model to detect referable active PDR by annotating new vessels (NV) and preretinal hemorrhages (PRH) (Fig. 1).

METHODS

Design and Setting

The development and validation of the DL system was conducted as a diagnostic accuracy study using retinal images from patients living on the island of Funen, Denmark. We retrospectively collected retinal images from patients with type 1 or 2 diabetes, which were obtained through the Danish screening program for DR [18].

Study Population

The study population was identified using The Funen Diabetes Database (FDDB) to which all screening episodes are reported at Odense University Hospital, Odense, Denmark [19]. The FDDB contains data of screening dates, DR severity levels, and demographic and clinical data. We included patients with PDR from 2009 to 2019 from a local image database. Data linkage for each individual is possible using the Civil Personal Register (CPR) number, a unique ten-digit number assigned to every child born in Denmark since year 1968.

Classification of Retinal Images

Figure 2 illustrates the process from the initial set of PDR images to the final dataset prepared for training and testing of the DL algorithm. Three medical doctors were all trained and certified by a Virtual Ocular Learning Platform (VIOLA) to become certified DR graders before establishing the reference standard in the present study [20]. The classification process started with grader 1 identifying and manually classifiying 3262 retinal images as active or inactive PDR from the entire population of patients with ICDR level 4. We defined a retinal image with active PDR if it exhibited at least one NV and/or at least one PRH. Out of the 3262 included ICDR level 4 images, no exclusion were made for images not meeting the ICDR level 4 criteria. Secondly, certified graders 2 and 3 each reassessed a randomly picked half of the images to ensure correct classification and ensure a high-quality reference standard. In cases of disagreement between grader 1 and either grader 2 or 3 on the classification of active versus inactive PDR, grader 1 reassessed the image with either grader 2 or 3 for a final decision. The dataset consisted of high-resolution 6528 × 6528 pixel six-field retinal images from a TopCon camera (Fig. 1) from which we included images of varying quality. Five images were excluded for not being gradable.

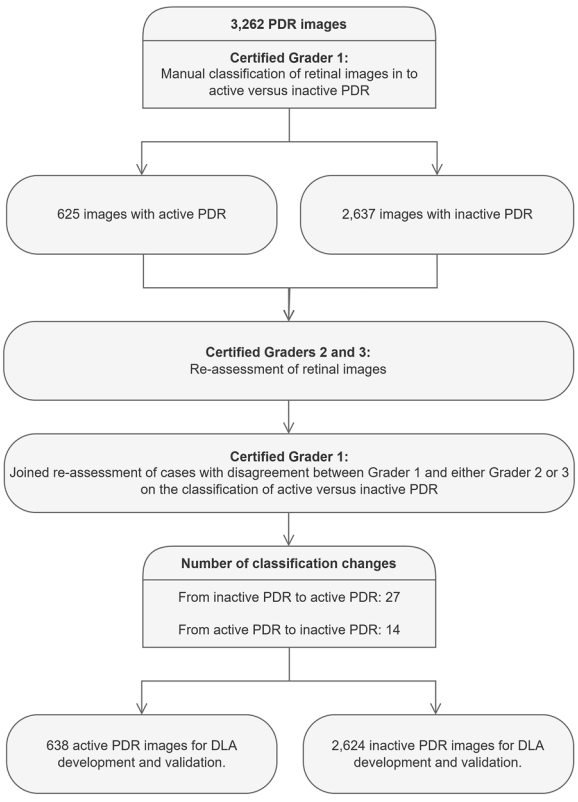


Fig. 2 Flowchart for the preparation of image data in the study. *PDR* proliferative diabetic retinopathy, *DLA* deep learning algorithm

Pre-segmentation of DR Lesions

Following the classification process, this step involved segmentation of DR lesions before training of the DL model. We applied our established DL segmentation system to the retinal images, which can annotate nine different DR lesions, including microaneurysms, soft exudates, hard exudates, intraretinal hemorrhages, intraretinal microvascular abnormalities (IRMA), NVs, PRHs and central as well as peripheral photocoagulation tracks [10].

A review of the pre-segmented retinal images then took place, which included assessing all segmentations of PRHs and NVs by reviewing and refining the automated visual drawings. Grader 1 evaluated the segmentations of all NVs and PRHs in the active PDR images, which included redrawing all of the pre-segmentations made by the DL system to ensure complete precision. Additionally, grader 1 identified and segmented any missed NV and PRH lesions that the DL system failed to annotate and deleted wrong segmentations. If the DL segmentation system incorrectly annotated NVs and PRHs in the images classified as inactive PDR, grader 2 and 3 manually corrected these errors. No images were reclassified after the pre-segmentation process.

We utilized the Computer Vision Annotation Tool (CVAT) version 2.3.0, which is an open-source software platform. The process began with uploading the retinal image data and the pre-annotations generated by the DL system. We zoomed in on the lesions to ensure optimal precision during the annotation of NVs and PRHs using the polygon segmentation tool within

CVAT. Finally, we extracted both the images and the refined annotation data from the CVAT platform ready for algorithm development.

Definition of Active PDR Detection

The DL model was developed as a segmentation model, meaning that it highlights DR lesions with visual overlays for clinicians to review. Rather than directly classifying active PDR, the DL model detects NVs and PRHs, which we then defined as a positive active PDR classification.

Outcome Measures

Primary outcome measures included the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the DL system to detect active PDR through segmentation of NVs and/or PRHs. Secondary outcome measures included an initial sensitivity and PPV for the DL system's ability to correctly detect NVs and PRHs on a lesion-by-lesion basis.

Algorithm Development

The active PDR images were split into training (70%), validation (10%), and testing (20%) datasets as shown in Table 1. We made sure that images from a unique individual only figured in one of the three datasets to prevent the DL overperforming on formerly seen data. A total of 301 inactive PDR images were additionally included in the testing dataset for validation of the diagnostic performance of the DL model. All images had detailed pixel-level annotations for NVs, PRHs, microaneurysms, hemorrhages, hard exudates, cotton wool spots, IRMA, photocoagulation scars, and the optic disc. The annotations for all images were produced by a previously developed segmentation model as described in Sect. "Pre-segmentation of DR Lesions" [10]. All annotations were reclassified into four categories to train the model for segmentation of NVs and PRHs: background, NV or PRH, and other. The "other" category included the aforementioned lesion types as well as noise arising from incorrect segmentations by the pre-segmentation model.

The DL segmentation model based on Unet++ [21] with an ImageNet pre-trained ConvNext backbone [22] was trained using 440 retinal images of active PDR with 1381 pixel-level annotations of NVs and PRHs, and tuned using 63 images with 123 NV and PRH annotations. The model was implemented in Pytorch using the TorchSeg library [23], and the model's output was generated using three convolution kernels of sizes 1×1, 3×3, and 5×5. Each kernel produced a separate probability map, which were then combined into a single prediction map used for loss computation.

We combined Log Cosh Dice Loss [24] and the Focal Tversky Loss [25] into a single loss function referred to as Focal Log Cosh Tversky Loss to address the challenge of imbalanced data. In this loss function, the Dice component of the Log Cosh Dice Loss was replaced with the Focal Tversky Loss using default weight and focusing parameters. Together, the Log Cosh component and Focal Tversky component creates a smooth optimization path that handles the class imbalance inherent in the data. We

Table 1 Numbers of images and segmentation counts for preretinal hemorrhages and new vessels across dataset splits

	Patients, *n*	Retinal images, *n*	Active PDR images, *n* (%)	New vessels and preretinal hemorrhages, *n* (%)
Training	105	440	440 (68.9)	1381 (73.5)
Validation	29	63	63 (9.9)	123 (6.5)
Testing	65	436	135 (21.2)	374 (19.9)
Total	199	939	638	1878

PDR proliferative diabetic retinopathy

split the images into patches of 288 × 288 pixels with an overlap of 20 pixels to avoid loss of information from the edges of the patches. Random 256 × 256 pixel crops were sampled from these in each iteration before training the model. Images were augmented with the albumentations library [26] using different spatial and pixel-level transforms. The labels were also augmented using random morphological erosion and dilation to handle label uncertainty and increase the robustness of the model.

In the beginning of the training process, we upsampled the patches with NVs or PRHs to match the number of patches with annotations for the other lesion types only. We randomly subsampled patches containing only background to match the same quantity at each epoch's start. For every tenth epoch, the fraction of background patches was increased until 100% of background patches were included. We optimized the model with the Adam optimizer using a cyclical learning rate schedule with learning rates ranging from 5×10^{-6} to 5×10^{-4}. The model was check pointed during training. We included the weights yielding the highest Dice coefficient score at the point after all background patches for validation on the testing dataset. Segmentations on the entire six-field validation images were produced using overlapping tiles and edge cropping.

The initial NV and PRH lesion detection test revealed a sensitivity of 100% and a PPV of 12% that resulted in a high number of false positive cases. We then conducted a qualitative review of the model segmentations and identified two types of errors compared to the reference standard. The qualitative review included scrolling through testing images with the DL segmentation masks (Fig. 4) to identify errors contributing to the low PPV. The first involved dot-like segmentations of very few pixels in random areas. The second type of error involved misinterpretation of irregular structures, such as IRMA. To address this, we adjusted the DL system at the image level to a sensitivity threshold of 0.9 causing the model to ignore lesions smaller than 10,576 pixels, which represented approximately 0.02% of the total area of a six-field retinal image.

Statistics

The DL model's performance metrics and 95% confidence intervals (95% CI) for sensitivity, specificity, PPV, and NPV were calculated using STATA17 (StataCorp, College Station, Texas, USA).

Approvals

We obtained approval for the study from the Danish Data Agency (journal number 22/9016) and the Regional Council of the Region of Southern Denmark (journal number 23/10823). The study adhered to the tenets of the Declaration of Helsinki. The register-based data used in this study were non-identifiable.

RESULTS

The study included 638 images with active PDR and 301 images with inactive PDR from 199 individuals. After the initial manual classification of active versus inactive PDR, the second classification process resulted in 27 images originally classified as inactive PDR and 14 images originally classified as active PDR being reclassified (Fig. 2). The prevalence of active PDR was 19.5% in the full dataset of 3262 images.

The training dataset consisted of 440 active PDR images with 1381 active PDR lesions from 105 patients, the validation dataset of 63 active PDR images with 123 active PDR lesions from 29 patients, and the testing dataset consisted of 135 active PDR images with 374 active PDR lesions and 301 inactive PDR images from 65 patients. The exact percentage distribution of images and lesions is provided in Table 1.

The confusion matrix in Fig. 3 shows that the DL model correctly identified 121 true positives and 211 true negatives from the total of 436 images in the testing dataset. It incorrectly classified 91 inactive PDR images as false positives and 13 active PDR images as false negatives.

The DL model's diagnostic performance at the image level for detection of active PDR reached a sensitivity of 90% (95% CI 85–95%),

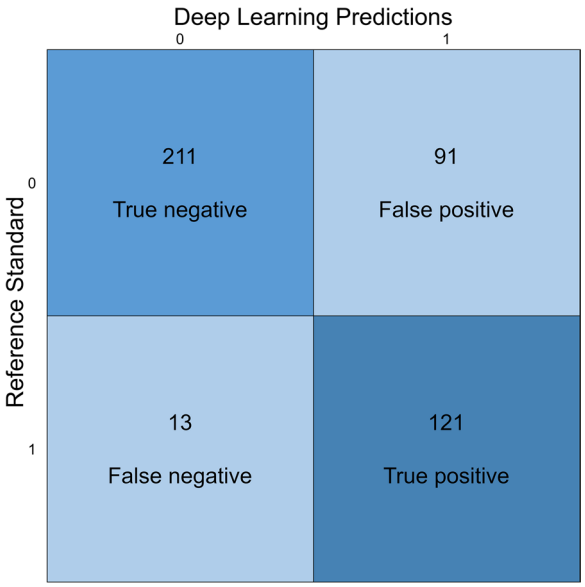


Fig. 3 Confusion matrix showing true positive, false positive, false negative, and true negative predictions by the deep learning model compared to the reference standard. The values represent absolute counts

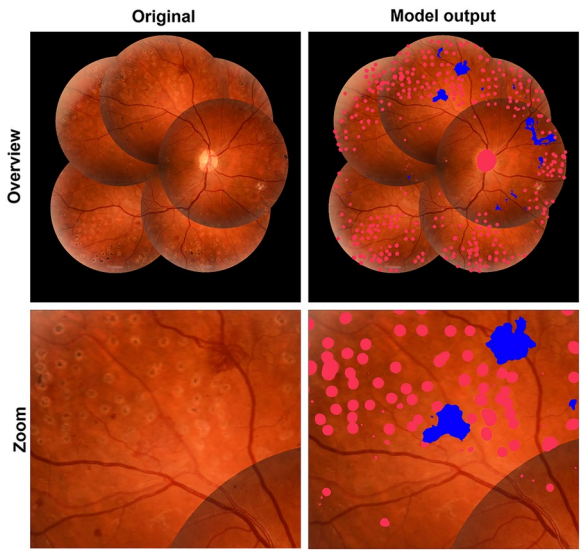


Fig. 4 An output image showing the annotations of new vessels by our deep learning segmentation model. The blue annotations are new vessels and all other annotations are in pink color

a specificity of 70% (95% CI 65–75%), a PPV of 57% (95% CI 50–64%), and an NPV of 94% (95% CI 91–97%) (Table 2). The DL model's output masks from a six-field image and a smaller region of the same image are illustrated in Fig. 4. The zoomed region demonstrates that while the model segments with reasonable precision, it also exhibits a tendency for dot-like segmentations, as described in Sect. "Algorithm Development". However, from the example in Fig. 4, the dot segmentation occurs within an NV, making it a true positive at the lesion level.

Table 2 Diagnostic performance of the deep learning model for detecting active PDR at the image level

	% (95% CI)
Sensitivity	90 (85–95)
Specificity	70 (65–75)
Positive predictive value	57 (50–64)
Negative predictive value	94 (91–97)

The definition of active PDR is the presence of one or more preretinal hemorrhages and/or new vessels. The table shows diagnostic accuracy at the 0.9 lesion threshold of 10,576 pixel (0.02% of the total image area). The reference is the established ground truth by the three certified graders

PDR proliferative diabetic retinopathy, *95% CI* 95% confidence interval

DISCUSSION

We developed a DL segmentation model with the ability to mark sight-threatening DR lesions and identify active PDR with a sensitivity of 90% and a negative predictive value of 94%. This feature not only enhances transparency but also improves the decision-making process for clinicians. We are encouraged by the sensitivity, which reached 90%, a critical threshold that ensures patient safety by reliably identifying those in need of further ophthalmological evaluation.

Early studies that investigated the DL classification of referable DR reported sensitivities of 87.2–97.5% and specificities of 85.3–93.4% [11–13, 15–17, 27]. However, given the inclusion of a very limited number of 16–70 images with

PDR [11, 27], these have not been developed to detect PDR nor distinguish between active and inactive disease. The studies utilized a variety of grading scales including the ICDR, the NHS guideline scale [28], and the Early Treatment Diabetic Retinopathy Study (ETDRS) scale [29]. They all applied the same threshold for referable DR corresponding to moderate non-proliferative DR (ICDR level 2). They also investigated the classification of vision-threatening DR with DL defined as severe non-proliferative DR (ICDR level 3) or above, achieving sensitivities from 95.2% to 100% and specificities from 89.5% to 97.4% [11, 12, 15, 16, 27]. In contrast, our study employs a more stringent threshold for both referable and vision-threatening DR by identifying those in need of immediate treatment.

Other DL systems have been developed to detect NVs [30–32], with two of them also incorporating segmentation models for NVs [31, 32]. The majority of the models are trained on relatively small datasets because of a general lack of retinal images with NVs. Small datasets increase the risk of overfitting and restrict a model's ability to generalize, as it may not be exposed to different morphologic variations of NVs and PRHs, which potentially impacts the performance in some clinical scenarios [33]. A study by Tang et al. [31] used 20 one-field retinal images and divided them into 20 patches per image. They reported an average sensitivity of 87.7% and a specificity of 99.7% for NV segmentation, while the sensitivity for classifying active PDR was 94.6% and the specificity 97.7%. In another study conducted by Setiawan et al. [30], features from pre-trained neural networks were extracted and used for classification of NVs in the optic disc area resulting in an accuracy of 90–100% from three different models. They used publicly available retinal images and developed their DL system restricted to 100 patches of the optic disc region with and without NVs. The study most comparable to ours was published by Alam et al. [32], which also pre-processed retinal images using a segmentation model before training a DL model for the classification of active PDR and the segmentation of NVs, but not PRHs. They included 1163 macula-centered 30° retinal images, of which 52% were non-PDR images. Of the remaining images, 418 were classified as active PDR and the last 141 images as inactive PDR. Their DL model achieved an accuracy of 87.7% in detecting active PDR in macula-centered images. These DL systems demonstrated high sensitivities and specificities for detecting vision-threatening DR. However, most of the studies rely on less stringent thresholds or were tailored to image modalities that do not align with the Danish clinical setting.

It was essential for us to develop a DL system tailored to the specific requirements of the national DR screening program of Denmark and, in general, a DL system with a clinical consequence when a patient is identified as screening positive (active PDR). This included aligning to the recommended threshold for referral to the hospital, with ethnic and systemic characteristics of the Danish population as well as the six-field retinal image modality.

It is widely believed that artificial intelligence can optimize resource allocations and reduce costs in healthcare. To fulfill this potential, DL models must first accurately identify patients in need of treatment. Secondly, to reduce healthcare costs, a model must also be trusted to identify healthy individuals. DL systems are shown to achieve sensitivity comparable to human expert graders for vision-threatening DR, but humans still excel in specificity [11]. Our DL model demonstrated sensitivity comparable to the 82.7–88.5% achieved by human expert graders for detecting vision-threatening DR, but exhibited a specificity of 70% largely due to false positive segmentations of NVs, which was low compared to human specificity for vision-threatening PDR of the 90.3–99.6% reported in two studies [11, 17]. It became evident that many of these false positives were caused by random and small dot-like segmentations of NVs. By removing small NV segmentations under 10,576 pixels from the analysis, we were able to achieve a specificity of 70%. Despite this improvement, our model still produces a fair number of false positives, primarily due to misinterpreting IRMA as NVs. The irregular shapes of both lesions pose a challenge for differentiation by the DL model, a task that can also be difficult for human retinal experts. Fluorescein angiography is sometimes required to confirm the diagnosis of NVs, which highlights the inherent difficulty of this task.

Improving specificity would require more image data, which is generally sparse for active PDR, and training of the model on inactive images might further enhance its positive predictive value.

A particular strength of our DL model lies in its integrated capability to provide visual segmentations of NVs and PRHs, which distinguishes it from many other models in the field of detecting vision-threatening DR [11–13, 16]. Unlike typical DL models that function as black boxes, our model not only classifies retinal images for vision-threatening DR but also leaves an argument for the clinician, and this feature could have an effect on the overall acceptance of artificial intelligence in clinical use. However, this segmentation, guided by initial human annotations, may constrain the model's potential to that of a human understanding, as it is trained to focus on predefined features of active PDR. Thus, while enhancing interpretability, our approach may inherently limit the ability of the DL model to discover novel patterns beyond current human knowledge such as contours of the optic disc, changes in retinal vascular diameter, tortuosity, and more. The present study is limited by a lack of external validation and also by the general lack of images representing active PDR. Although we only train the model on cases, NVs and PRHs examples remain underrepresented, which necessitates the use of upsampling techniques. However, further upsampling could lead to an increased risk of overfitting, as the model may become too specialized to the training set examples. We assume that the background and other lesion examples in the images are sufficiently representative for the model to learn to differentiate effectively. Future research will need focus on external testing of the developed DL algorithm to ensure its generalizability.

CONCLUSION

The DL model demonstrated an excellent sensitivity (90%) and an acceptable specificity (70%) for detecting active PDR, meeting the stringent requirements of the Danish DR screening program. The model provides insight into its decision-making and enhances clinical decision-making of patients requiring urgent treatment by providing lesion-level annotations of NVs and PRHs.

Medical Writing. No writing assistance beside inputs from co-authors was received during the writing of this article.

Author Contributions. Sebastian Dinesen: Concept, study design, image preparation, drafting manuscript; Marianne Guldager Schou: Image data preparation, drafting manuscript; Christoffer Viuff Hedegaard: Image data preparation, drafting manuscript; Yousif Subhi: Drafting manuscript; Thiusius Rajeeth Savarimuthu: Concept, study design, algorithm development; Tunde Peto: Concept, study design, drafting manuscript; Jakob Holm Andersen: Algorithm development, drafting manuscript; Jakob Grauslund: Concept, study design, drafting manuscript. All named authors fulfilled the criteria for co-authorship according to the International Committee of Medical Journal Editors (ICMJE). All the authors take responsibility for the work and gave their approval for this version to be published.

Funding. The study received PhD salary contributions from the Steno Diabetes Centre Odense (application 4836), Odense, Denmark and The Region of Southern Denmark (J.nr. 21/18139), Vejle, Denmark. Funding for rapid service fees was provided by Overlægerådets Forskningsfond, Odense University Hospital, Odense, Denmark. The funding organizations had no role in the design or conduct of this research.

Data Availability. The datasets generated and analyzed during the current study are not publicly available due to General Data Protection Regulations (GDPR). Anyone can apply for access to data in The Funen Diabetes Database.

Declarations

Conflict of Interest. Author Jakob Grauslund declares to have received speaker's fee from Allergan, Bayer, Novartis, and Roche, and to have served as an advisory board member for Allergan, Apellis, Bayer, Novartis, and Roche, not related to this work. Author Yousif Subhi declares to have received speakers' fee from Bayer and Roche and to be the inventor of a patent related to biomarkers for polypoidal choroidal vasculopathy. Yousif Subhi is an Editorial Board member of *Ophthalmology and Therapy*. Yousif Subhi was not involved in the selection of peer reviewers for the manuscript nor any of the subsequent editorial decisions. (WO2020007612A1), not related to this work. Sebastian Dinesen, Tunde Peto, Marianne G. Schou, Christoffer V. Hedegaard, Thiusius R. Savarimuthu and Jakob K. H. Andersen declare that no conflicts of interests exist in relation to this work.

Ethical Approvals. We obtained approval for the study from the Danish Data Agency (journal number 22/9016) and the Regional Council of the Region of Southern Denmark (journal number: 23/10823). The study adhered to the tenets of the Declaration of Helsinki. The register-based data used in this study were non-identifiable.

REFERENCES

1. GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet. 2023;402(10397):203–34.
2. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. Ophthalmology. 2021;128(11):1580–91.
3. Antonetti DA, Klein R, Gardner TW. Diabetic retinopathy. N Engl J Med. 2012;366(13):1227–39.
4. Yau JW, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care. 2012;35(3):556–64.
5. Abou Taha A, Dinesen S, Vergmann AS, Grauslund J. Present and future screening programs for diabetic retinopathy: a narrative review. Int J Retina Vitreous. 2024;10(1):14.
6. Grauslund J, Stokholm L, Andersen N, et al. Risk of 2-year progression from no diabetic retinopathy to proliferative diabetic retinopathy in accordance with glycaemic regulation in 145,527 persons with diabetes in a national cohort. Acta Ophthalmol. 2024;102(3):e407–9.
7. Balyen L, Peto T. Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. Asia Pac J Ophthalmol (Phila). 2019;8(3):264–72.
8. Nakayama LF, Zago Ribeiro L, Novaes F, et al. Artificial intelligence for telemedicine diabetic retinopathy screening: a review. Ann Med. 2023;55(2):2258149.
9. Nielsen KB, Lautrup ML, Andersen JKH, Savarimuthu TR, Grauslund J. Deep learning-based algorithms in screening of diabetic retinopathy: a systematic review of diagnostic performance. Ophthalmol Retina. 2019;3(4):294–304.
10. Andersen JKH, Hubel MS, Rasmussen ML, Grauslund J, Savarimuthu TR. Automatic detection of abnormalities and grading of diabetic retinopathy in 6-field retinal images: integration of segmentation into classification. Transl Vis Sci Technol. 2022;11(6):19.
11. Ting DSW, Cheung CY, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA. 2017;318(22):2211–23.

12. Chia MA, Hersch F, Sayres R, et al. Validation of a deep learning system for the detection of diabetic retinopathy in Indigenous Australians. Br J Ophthalmol. 2024;108(2):268–73.

13. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316(22):2402–10.

14. Raumviboonsuk P, Krause J, Chotcomwongse P, et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. NPJ Digit Med. 2019;2:25.

15. Ipp E, Liljenquist D, Bode B, et al. Pivotal evaluation of an artificial intelligence system for autonomous detection of referrable and vision-threatening diabetic retinopathy. JAMA Netw Open. 2021;4(11):e2134254.

16. Li Z, Keel S, Liu C, et al. An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. Diabetes Care. 2018;41(12):2509–16.

17. Dai L, Wu L, Li H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun. 2021;12(1):3242.

18. Grauslund J, Andersen N, Andresen J, et al. Evidence-based Danish guidelines for screening of diabetic retinopathy. Acta Ophthalmol. 2018;96(8):763–9.

19. Adelborg K, Szentkúti P, Henriksen JE, et al. Cohort profile: the Funen Diabetes Database-a population-based cohort of patients with diabetes in Denmark. BMJ Open. 2020;10(4):e035492.

20. Andersen JKH, Hubel MS, Savarimuthu TR, Rasmussen ML, Sørensen SLB, Grauslund J. A digital online platform for education and certification of diabetic retinopathy health care professionals in the Region of Southern Denmark. Acta Ophthalmol. 2022;100(5):589–95.

21. Ronneberger O, Fischer P, Brox T, editors. U-Net: convolutional networks for biomedical image segmentation. Cham: Springer; 2015.

22. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): IEEE Computer Society; 2016. p. 2818–26.

23. PyTorch. https://github.com/isaaccorley/torchseg. July 2024.

24. Jadon S, editor. A survey of loss functions for semantic segmentation. In: 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); 27–29 Oct 2020.

25. Abraham N, Khan NM, editors. A novel focal tversky loss function with improved attention U-net for lesion segmentation. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019); 8–11 April 2019.

26. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: fast and flexible image augmentations. Information. 2020;11(2):125.

27. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med. 2018;1:39.

28. Scanlon PH. The English National Screening Programme for diabetic retinopathy 2003–2016. Acta Diabetol. 2017;54(6):515–25.

29. Early Treatment Diabetic Retinopathy Study Research Group. Early photocoagulation for diabetic retinopathy. ETDRS report number 9. Ophthalmology. 1991;98(5):766–85.

30. Setiawan W, Mohammad Imam U, Rulaningtyas R. Classification of neovascularization using convolutional neural network model. Telkomnika (Telecommunication Computing Electronics and Control). 2019;17:463–72.

31. Tang MCS, Teoh SS, Ibrahim H, Embong Z. Neovascularization detection and localization in fundus images using deep learning. Sensors (Basel). 2021;21(16):5327.

32. Alam M, Zhao EJ, Lam CK, Rubin DL. Segmentation-assisted fully convolutional neural network enhances deep learning performance to identify proliferative diabetic retinopathy. J Clin Med. 2023;12(1):385.

33. Rather IH, Kumar S, Gandomi AH. Breaking the data barrier: a review of deep learning techniques for democratizing AI with small datasets. Artif Intell Rev. 2024;57(9):226.