






Article

Artificial Intelligence-Based Screening System for Diabetic Retinopathy in Primary Care

Marc Baget-Bernaldiz ^{1,*}, Benilde Fontoba-Poveda ², Pedro Romero-Aroca ¹, Raul Navarro-Gil ¹,
Adriana Hernando-Comerma ¹, Angel Bautista-Perez ¹, Monica Llagostera-Serra ¹, Cristian Morente-Lorenzo ¹,
Montse Vizcarro ¹ and Alejandra Mira-Puerto ¹

- ¹ Ophthalmology Service, Hospital Universitari Sant Joan, Institut d'Investigació Sanitària Pere Virgili [IISPV], Universitat Rovira i Virgili, 43204 Reus, Spain; romeropere@gmail.com (P.R.-A.); raul_navarro_gil@hotmail.com (R.N.-G.); adriana.hernando.comerma@gmail.com (A.H.-C.); abautistaper@gmail.com (A.B.-P.); monillagostera@hotmail.com (M.L.-S.); cristian_ml87@outlook.com (C.M.-L.); montseviz@hotmail.com (M.V.); a.mira.puerto@gmail.com (A.M.-P.)
- ² Responsible for Diabetic Retinopathy Eye Screening Program in Primary Care in Baix Llobregat Barcelona (Spain), Institut d'Investigació Sanitària Pere Virgili [IISPV], 43204 Reus, Spain; bfontoba@gmail.com
- * Correspondence: mbaget@gmail.com

Abstract: Background: This study aimed to test an artificial intelligence-based reading system (AIRS) capable of reading retinographies of type 2 diabetic (T2DM) patients and a predictive algorithm (DRPA) that predicts the risk of each patient with T2DM of developing diabetic retinopathy (DR). Methods: We tested the ability of the AIRS to read and classify 15,297 retinal photographs from our database of diabetics and 1200 retinal images taken with Messidor-2 into the different DR categories. We tested the DRPA in a sample of 40,129 T2DM patients. The results obtained by the AIRS and the DRPA were then compared with those provided by four retina specialists regarding sensitivity (S), specificity (SP), positive predictive value (PPV), negative predictive value (NPV), accuracy (ACC), and area under the curve (AUC). Results: The results of testing the AIRS for identifying referral DR (RDR) in our database were ACC = 98.6, S = 96.7, SP = 99.8, PPV = 99.0, NPV = 98.0, and AUC = 0.958, and in Messidor-2 were ACC = 96.78%, S = 94.64%, SP = 99.14%, PPV = 90.54%, NPV = 99.53%, and AUC = 0.918. The results of our DRPA when predicting the presence of any type of DR were ACC = 0.97, S = 0.89, SP = 0.98, PPV = 0.79, NPV = 0.98, and AUC = 0.92. Conclusions: The AIRS performed well when reading and classifying the retinographies of T2DM patients with RDR. The DRPA performed well in predicting the absence of DR based on some clinical variables.

Keywords: diabetic retinopathy; artificial intelligence; algorithm; diabetic retinopathy screening; primary care



Citation: Baget-Bernaldiz, M.; Fontoba-Poveda, B.; Romero-Aroca, P.; Navarro-Gil, R.; Hernando-Comerma, A.; Bautista-Perez, A.; Llagostera-Serra, M.; Morente-Lorenzo, C.; Vizcarro, M.; Mira-Puerto, A. Artificial Intelligence-Based Screening System for Diabetic Retinopathy in Primary Care. *Diagnostics* **2024**, *14*, 1992. <https://doi.org/10.3390/diagnostics14171992>

Academic Editors: Mobeen Ur Rehman and Kil-To Chong

Received: 13 August 2024
Revised: 1 September 2024
Accepted: 6 September 2024
Published: 9 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Type 2 diabetes mellitus (T2DM) is a chronic disease associated with insulin resistance that causes hyperglycemia. It mainly affects the brain, heart, kidneys, and eyes through vascular involvement, causing strokes, myocardial infarctions, nephropathy, and diabetic retinopathy (DR).

According to data from the International Diabetes Federation (IDF), the estimated global prevalence of T2DM in 2021 was 536 million people and is expected to affect 783 million people by 2045. However, the prevalence of diabetes varies widely across regions of the world. The prevalence is highest in the Middle East and North Africa, affecting 16.2% of their populations. As for North America and the Caribbean, the prevalence is 14%. On the other hand, Europe and Southeast Asia have the lowest prevalence, estimated at 9.2% and 8.7%, respectively [1]. In Spain, a study conducted by di@bet.es reported that 7.9%

of the population is known to have DM2, although its true prevalence is estimated to be double that, at 13.8% [2].

According to IDF, the global prevalence of diabetic retinopathy (DR) was 22.27% in 2020. This means that the number of adults worldwide with DR is 103.12 million and is expected to increase to 160.50 million by 2045. The prevalence of DR is highest in Africa (35.90%), North America (35%), and the Caribbean (33.30%) and is lowest in South and Central America (13.37%) [3].

Diabetic retinopathy is the leading cause of preventable vision loss and blindness in the working-age population, resulting in a major impact in terms of health and economy. Direct costs associated with DR increase as the disease stages progress, with costs being highest in the sight-threatening stages of DR (STDR). In the Singapore study by Zhang et al. [4], the total cost for patients with STDR was 3.79 times that of patients without DR. Similarly, a USA study showed that diabetic patients with STDR had twice the DR-related cost of the milder cases [4]. Finally, a study in Germany found that the ratio of DR-related costs for mild, moderate, and STDR relative to no DR were 1.2, 2.5, and 3.2, respectively [5].

All diabetic patients have the potential to develop DR; therefore, regular screening and early diagnosis are essential and recommended by ophthalmology associations. It can be detected by fundus examination or by retinography, which is more commonly recommended and forms the basis of current DR screening protocols [6,7].

In developed countries, DR screening is most often performed by ophthalmologists, who examine the patients' retinas under pupil dilation, or through telemedicine, where patients come to the screening center to have a photograph of their retinas taken, which is then read by experts remotely. That said, early detection of DR is difficult in the current screening systems because of such huge demand. In addition, a significant number of diabetic patients who live far away from screening sites fail to attend scheduled visits, and, as a result, fewer than 50% of diabetes patients have regular examinations [8].

Despite the effectiveness of teleretinal screening programs, they are not cost-effective because the manual grading requires many specialists to be involved in it [9]. Therefore, an inexpensive, accurate, and automated method for classifying fundus photographs for DR detection in the primary care clinical setting would greatly benefit providers, healthcare systems, and patients.

Over the last decade, several groups have developed artificial intelligence reading systems (AIRSs) capable of reading retinographies, aimed to alleviate the screening for DR. Most AIRSs are deep learning algorithms (DLAs) that provide a binary classification to distinguish patients with more evolved DR types or referable DR (RDR) from those with no or mild DR [10,11]. Currently, existing DLAs have demonstrated similar or even better performance than human experts on several DR classification tasks [10]. Furthermore, these AIRSs appear to be much more cost-effective than manual assessment, which is the current protocol. A study in Scotland showed that the costs of a national DR screening program can be halved by replacing manual grading with automated assessment [12]. Similarly, another UK study reported a cost saving of 21% [13].

To complement that, other research groups have developed DR prediction algorithms (DRPAs) based on clinical variables aimed at evaluating the risk to each diabetic patient of developing sight-threatening DR (STDR) and, thus, personalizing the screening schedules.

Across the world, primary care center teams are generally not involved in DR screening due to their lack of training and lack of equipment available, but in countries where primary care is better funded, most diabetic patients will visit their primary care center at some time each year for various reasons [14]. Health managers should try to take advantage of that by devising a screening system in which primary care teams are front and center.

In this article, we tested an AIRS aimed at identifying T2DM patients with RDR (moderate and severe DR) and a DRPA that can estimate the risk of developing DR to each T2DM patient according to nine clinical variables.

2. Overview of the Published Articles

Below is a summary of the most relevant published AIRSs and DRPAs. Regarding AIRSs, the IDX-DR (Digital Diagnostics, Coralville, IA, USA) was the first algorithm to obtain FDA approval in 2018 [15]. It obtained a sensitivity (S) of 96.8% and a specificity (SP) of 59.4% when detecting RDR. It was tested with Messidor-2, but the datasets used for training and validation were not provided.

Another AIRS that has obtained CE Class IIa Mark and FDA approval was EyeArt[®] (Eyenuk, Woodland Hills, CA, USA), which was independently validated and tested by the UK National Health Service (NHS) [16]. The results showed an S of 91.3% and an SP of 91.1% when detecting DR [17]. EyeArt was tested offline in India using the Remidio Fundus on Phone device to capture fundus images after pupillary dilation. This study achieved remarkable levels of S and SP for the detection of any degree of DR.

Retmarker[®] (retmarker SA, Taveiro, Portugal) also obtained CE Mark Class IIa approval. It was evaluated using over 100,000 fundus images from 20,258 consecutive NHS patients in the UK. It gave 95.8% S and 63.2% SP when detecting RDR. The datasets used for training and validation were not provided. An interesting feature of Retmarker is its ability to compare images to detect DR progression based on the number of microaneurysms present [18]. Furthermore, Retmarker demonstrated a 48% reduction in workload in DR detection in Portugal.

RetCAD (Thirona, The Netherlands) is an algorithm trained to detect DR and age macular degeneration. It achieved an S of 86.1% and an SP of 91.6% for RDR disease. It was validated using the MESSIDOR database [19].

The Eye Research Institute and the National University of Singapore developed SELENA+ to detect RDR, glaucoma, and age macular degeneration [20]. They designed the system to operate autonomously or semiautonomously, with human assistance. It was trained and tested on approximately 500,000 retinal images from diverse datasets. In 2019, SELENA+ was approved by the Singapore Health Service Authority to be implemented in the national DR screening program. Compared to 17 human assessors, SELENA+ performed with comparable accuracy and spent significantly less time.

In China, an AIRS was developed using over 70,000 retinographies from a web-based platform, LabelMe (Guangzhou, China). External validation of this algorithm was performed using over 35,000 images from population-based cohorts of Malaysians, Caucasian Australians, and Indigenous Australians. The S and SP were found to be 92.5% and 98.5%, respectively [21].

ARDA (Automated Retinal Disease Assessment) is an AIRS developed by Verily Life Sciences LLC (South San Francisco, CA, USA) [22]. This algorithm was developed using datasets of approximately 130,000 retinal photographs of patients with DM from the USA, and for the validation process, it used approximately 10,000 retinographies extracted from Messidor-2 and EyePACS datasets. The algorithm demonstrated an S of 96.8%, which was higher than that of the human graders (74%), while the SP was comparable (96–97%).

Finally, a retrospective validation study compared seven different AIRSs for detecting RDR that had been previously validated [23]. The investigators found that most of the algorithms performed no better than human graders. The sensitivities varied widely (51.0–85.9%), although high negative predictive values (82.7–93.7%) were observed. Interestingly, one algorithm was significantly worse than human graders, missing up to a quarter of cases of advanced DR (sensitivity of 72.4% for RDR), a limitation that could potentially lead to vision loss. The characteristics of the main validated AIRSs are shown in Table 1 together with our MIRA algorithm.

Table 1. Main validated AI-based algorithms for diabetic retinopathy detection.

Algorithm	Tech	Year	Training Set	Validation Set	Testing Set	DR Classification	Detection	AUC	S (%)	SP (%)
IDX-DR	CNN	2018	N/A	N/A	Messidor-2	FPRC	RDR	N/A	96.8	59.4
EyeArt 2.0	Image analysis technology	2019	EyePACS	N/A	850,908	ICDR	RDR	0.96	91.3	91.1
Retmarker	Recognition of lesions	2011	N/A	N/A	21,514	No RD/RDR	DR/no DR	0.84	95.8	63.2
SELENA+	CNN	2019	ImageNet	N/A	3556	ICDR	RDR	0.95	91.8	98.7
LabelMe	CNN	N/A	71,043	35,201	N/A	N/A	RDR	0.95	92.5	98.5
ARDA	CNN	N/A	130,000	Messidor-2 EyePACS	N/A	N/A	RDR	0.99	87	98.5
MIRA	CNN	2019	Own dataset EyePACS	Own dataset 5000	Own data Messidor-2	Messidor-2	RDR	0.92–0.958	94.6–96.7	99.1–99.8

AUC: area under the curve; S: sensitivity; SP: specificity; CNN: convolutional neural network; N/A: not available; FPRC: Wisconsin Fundus Photograph Reading Center; RDR: referral diabetic retinopathy; ICS; ICDR: International Clinical Diabetic Retinopathy Severity Scale. Any-DR: any type of RD. Our algorithm (MIRA) performed well in detecting RDR compared to the others.

Regarding DRPAs, three main algorithms have been developed that predict the presence of vision-threatening DR. Aspelund et al. [24] based their method on the level of HbA1c, type of diabetes, value of systolic pressure, the current age, and diabetes duration. It was constructed from the diabetic retinopathy database managed at the Ophthalmology Department of Aarhus University Hospital in Denmark. The database was built based on the clinical data of 5199 patients over 20 years, thus allowing the algorithm to be tested prospectively. The model provides a recommended interval for follow-up fundus screening for the presence of sight-threatening vision of between 6 and 60 months. The algorithm has been further tested in other countries, such as Spain, where Soto Pedre [25] studied 508 patients with T2DM. The results showed that 3.1% developed STDR before their subsequent screening visit, with the value of the area under the curve (AUC) being 0.74. Finally, in the UK, Lund [26], using a sample of 9690 DM patients followed for 2 years, reported that the algorithm predicted the onset of DR stages with AUC values of 0.833 for the T2DM patients.

The Scanlon method is based on the age of patients and their levels of HbA1c and cholesterol [27]. This algorithm was validated with 15,877 patients, obtaining an AUC of 0.77 for predicting the development of STDR.

Finally, the model developed by Broadbent [28], also known as the Liverpool Risk Calculation Engine, was built to detect the risk of developing STDR. The statistical analysis of this algorithm found an AUC value of 0.88 in the prediction of STDR, with an S of 0.61 and an SP of 0.93. We should highlight that this is the only report that provided sensitivity and specificity data. This algorithm was tested on a population of DM patients from Liverpool (UK) and was published in the form of an action protocol called the ISDR protocol. Tables 2 and 3 describe the clinical variables used to construct each DRPA and their performance compared to our algorithm (Retiprogram), respectively.

Table 2. Algorithms for predicting the presence of diabetic retinopathy.

Algorithm	Age	Sex	DM Duration	DM Treatment	HbA1c	HTA	eGFR	Protein in Urine	BMI	Type DM	Cholesterol
Aspelund	X		X		X	X				X	
Scanlon	X				X						X
Broadbent					X						X
Retiprogram	X	X	X	X	X	X	X	X	X		

Clinical variables used to build each algorithm. Our algorithm (Retiprogram) is based on 9 clinical variables.

Table 3. Results obtained by the different diabetic retinopathy predicting algorithms in the testing phase.

Author	Testing Place (Group)	Number of Patients	AUC
Aspelund	Spain (Soto-Pedre)	508	0.74
	UK (Lund)	9690	0.83
Scanlon	Ireland (Smith)	2929	0.77
Broadbent	UK	4460	0.88
Retiprogram	Catalonia (Romero)	40,129	0.92–0.96

Our DRPA (Retiprogram) performed very well compared to the others.

3. Subjects

3.1. Setting

In the database of the Catalan public health system (SIDIAP), there are about 600,000 patients registered with DM2. For 265,388 of these cases, we had access to their electronic health records (EHRs) and their retinographies. On the other hand, in our database (Hospital Universitari Sant Joan de Reus, Catalonia, Spain), there are currently 21,087 T2DM patients registered, for whom we have their complete EHR and retinographies.

3.2. Datasets

To test the AIRS, we used 17,297 fundus images, of which 16,097 corresponded to 7389 patients in our database who had undergone annual eye screening for DR during the period from 1 February 2017 to 8 September 2023, and the 1200 fundus images that comprise the public Messidor-2 database [29]. Previously, we had used two different retinographies datasets to build and train the AIRS, with a total of 103,815 images: our sample of 15,123 tagged fundus images (different patients from the validation and testing phases) and 88,692 retinal images extracted from EyePACS [30]. Finally, 5000 fundus images from our database were used to validate the AIRS.

We used the EHR of 40,129 patients with type 2 diabetes mellitus from the SIDIAP database for whom we had the nine clinical variables necessary to test the DRPA during the 11-year follow-up period. We had previously used the EHR of a sample of 101,802 patients with type 2 diabetes mellitus extracted from SIDIAP to validate the predictive algorithm.

3.3. Design

To test the ability of the AIRS to read and classify fundus images, we compared the DR categories obtained by the algorithm with those DR categories assigned by a panel of retina experts. In addition, to test the performance of the DRPA, we compared the binary classification predicted by the algorithm (presence or absence of any type of DR) with the classification provided by retina specialists.

3.4. Inclusion and Exclusion Criteria

3.4.1. Inclusion Criteria

- Type 2 diabetic patients included in the SIDIAP database for whom we had all the clinical and demographic variables necessary to feed the DRPA during the 11-year follow up study period.
- Type 2 diabetic patients from our database and Messidor-2 database with high-quality retinographies to feed the AIRS.

3.4.2. Exclusion Criteria

- Type 1 diabetic patients.
- Gestational diabetes.
- Patients who did not give informed consent.
- Type 2 diabetic patients with incomplete EHR or poor-quality retinographies.

4. Materials and Methods

4.1. Ethics and Consent

CEIM IISPV (Institut d'Investigació Sanitària Pere Virgili) approved the present study (approval code RetinaReadRisk, protocol version 1. 3 October 2022, Reference number CEIM: 071/2022), which was carried out in accordance with the revised guidelines of the Declaration of Helsinki. All patients included in the study were previously informed about its objective. Once their commitment to participate in the follow-up of the study was obtained, they were asked to sign the informed consent.

4.2. The Algorithms

4.2.1. Artificial-Intelligence-Based Reading System

Model Construction and Training

Briefly, our baseline model used a $3 \times 640 \times 640$ input image obtained from a minimal preprocessing step, in which only the external background borders were trimmed and later resized to the required input size. The model consists of a convolutional neural network with 7 blocks of 2 layers each that progressively reduce the size of the data until it has a receptive field of $64 \times 5 \times 5$ for feature extraction. Each layer is a stack of a 3×3 convolution with stride 1×1 and padding 1×1 , followed by batch normalization and an ReLU activation function. The final vector has a size of 64 values, which is obtained from a 4×4 average pooling stage. In the last layer, a linear classifier and a softmax function use these 64 features to calculate the probability of each of the DR levels. For optimization of the parameters of this convolutional neuronal network, the quadratic weighted kappa is used as a loss function, because it is more appropriate for ordinal classification [13]. Details of the architecture can be found in this work [31].

All 88,692 retinographies available from EyePACS were used [24] together with a sample of 15,123 tagged retinographies from our database to build and train the AIRS. There were 81,266 retinal images with no DR, 8771 with mild DR, 14,097 with moderate DR, and 4588 with severe or proliferative DR in the training phase. As the incidence of proliferative DR was low in both datasets, we decided to merge the severe and proliferative into the same category [32].

Validation

A total of 5000 retinographies were taken from our DR screening database for validation. The results of the DLA when detecting RDR were sensitivity = 0.998, specificity = 0.958, positive predictive value = 0.711, negative predictive value = 0.929, error type I = 0.032, and error type II = 0.001 [32]. The AIRS MIRA 4.0 software was granted registration code SAFE CREATIVE 2007104712196 on 10 July 2020 at 11:24 UTC.

Testing

For testing, we used 15,297 retinographies from our database and all 1200 retinal images from Messidor-2. First, the reading was made by the AIRS, and then a second reading was made by four masked senior retina ophthalmologists. Finally, the agreement between ophthalmologists and the AIRS for detecting RDR was calculated.

Diabetes Retinopathy Classification

Both retina specialists and the AIRS used the Messidor-2 classification for grading the retinographies into levels of severity. The presence of fewer than 5 microaneurysms on a retinogram was classified as mild DR. When there were more than 5 but fewer than 15 microaneurysms or fewer than 5 retinal hemorrhages, it was classified as moderate DR. Finally, when there were more than 5 retinal hemorrhages or the presence of new vessels, it was classified as severe DR (Table 4) [29].

Table 4. Diabetic retinopathy (DR) classification according to the criteria used in the Messidor-2 database.

DR Grade	Microaneurysms (μ A)	Hemorrhages (H)	Neovascularization
No DR	0	0	0
Mild DR	≤ 5	0	0
Moderate DR	$5 < \mu A < 15$	$0 < H < 5$	0
Severe DR	$\mu A \geq 15$	$H \geq 5$	0/1

μ A: microaneurysms; H: hemorrhages; NV; 0 = absence of neovascularization; 1 = presence of neovascularization.

4.2.2. Diabetic Retinopathy Prediction Algorithm

Model Construction and Training

The DRPA consists of a fuzzy random forest of 100 decision trees and three independent variables in each node. To develop the DRPA, we initially used a sample of 2323 type 2 diabetic patients from our database of whom we had all the medical variables to feed it [33]. Then, we retrained the algorithm with a much bigger sample of 139,658 patients extracted from SIDIAP (System for Research and Development in Primary). The output from the DRPA predicted a binary result, the presence or absence of DR. First, we included 18 variables (current age, age at diagnosis of type 2 DM, sex, body mass index (BMI), diabetes duration, diabetes treatment, smoker status, blood pressure control, diastolic tension rate, systolic tension rate, HbA1c%, creatinine, estimated glomerular filtration rate (eGFR), total cholesterol, LDL-cholesterol, HDL-cholesterol, triglycerides, and microalbuminuria. By statistical analysis, we evaluated these variables, and only 9 results were significant after applying the fuzzy random forest model. Finally, we decided to build the DRPA using these 9 variables: current age, sex, diabetes duration, diabetes treatment, good or bad control of blood pressure (bad control was defined as systolic arterial tension > 140 mm Hg or diastolic arterial tension > 90 mm Hg), HbA1c level, eGFR, the microalbuminuria value, and the BMI.

It estimates the risk to each diabetic patient of having any type of DR according to 9 clinical variables, allowing us to personalize the screening time interval for each patient (Figure 1). Further details are given in our previously published work [34].

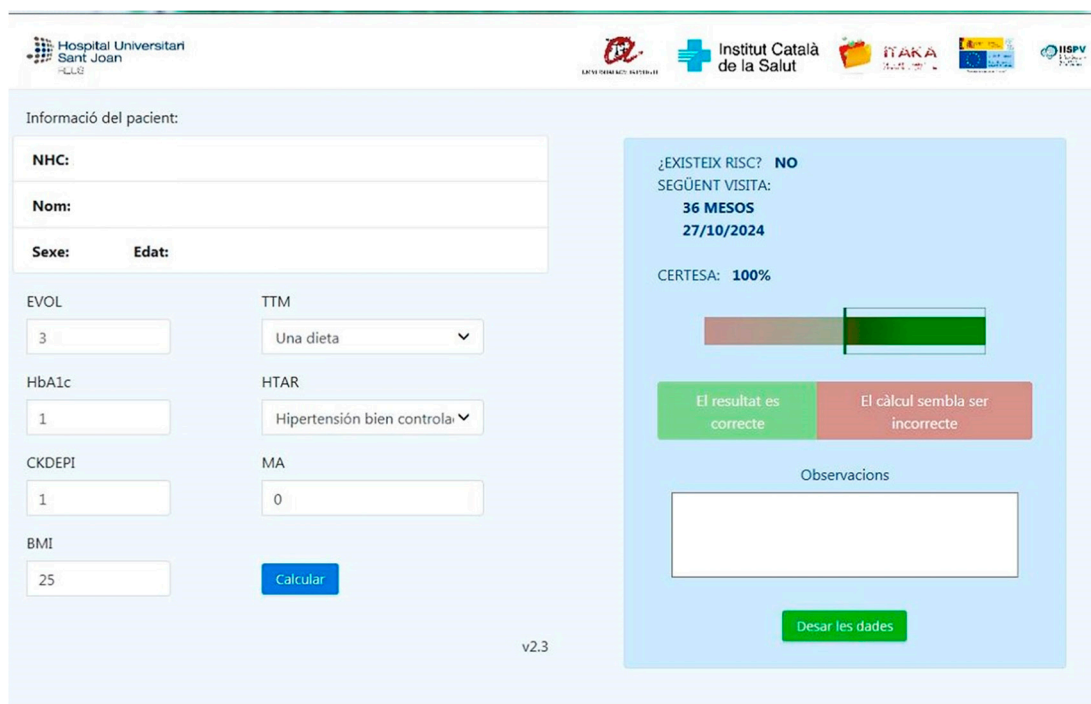


Figure 1. Diabetic retinopathy predicting algorithm interface embedded in the Catalan Health System. In this example, our algorithm (Retiprogram) considered this diabetic patient to be at low risk of developing diabetic retinopathy based on clinical variables and extended the next visit to 36 months.

Validation

Validation was made using the EHR of a sample of 101,802 T2DM patients taken from SIDIAP. Table 5 shows the clinical data of the population with and without DR used to validate the DRPA. The prevalence of DR in the sample was 19,759 patients (19.9%). Microalbuminuria was present in 16,196 patients (14.9%), nephropathy in 1650 patients (1.5%), and dyslipidemia in 26,994 patients (24.9%). There were differences between groups regarding diabetes duration, BMI, glycosylated hemoglobin, microalbuminuria, estimated glomerular filtration rate (eGFR), and arterial hypertension [34].

Table 5. Clinical and demographic variables used to validate the predictive algorithm for the presence of diabetic retinopathy.

Variable	Without DR	With DR	<i>p</i>
Age (y)	68.53 ± 11.06 (30–99)	69.89 ± 9.89 (33–98)	0.684 ¹
Female (%)	46.67	48.40	0.3802 ²
Diabetes duration (y)	7.25 ± 5.20 (0.2–56.99)	11.15 ± 6.90 (0.2–48.87)	<0.001 ¹
HbA1c (%)	7.22 ± 1.26 (3.5–16.5)	7.82 ± 1.45 (3.8–18.50)	<0.001 ¹
Microalbuminuria (mg)	34.73 ± 132.65 (0–59.75)	81.08 ± 250.74 (16.24–2999.75)	<0.001 ¹
Body mass index	30.21 ± 5 (16–38.91)	30.15 ± 5.15 (16.24–40.75)	0.004 ¹
Creatinine	1.13 ± 0.23 (0.87–1.22)	1.16 ± 0.35 (0.87–1.23)	<0.001 ¹

Table 5. Cont.

Variable	Without DR	With DR	<i>p</i>
eGFR (CKD-EPI)	60.62 ± 7.56 (60.05–69.84)	58.55 ± 9.54 (58.53–69.77)	<0.001 ¹
Arterial hypertension (%)	33	39	<0.001 ²
Cholesterol total	196 ± 41.3 (165–258)	198 ± 43.4 (168–261)	0.883 ¹
Triglycerides	168 ± 122 (42–298)	168 ± 125 (40–301)	0.386 ¹

¹ Chi-squared test; ² Student's *t*-test. The DR group showed a longer diabetes duration and a higher proportion of patients with arterial hypertension. They also showed higher values of HbA1c, microalbuminuria, BMI, creatinine, and eGFR compared to the others.

The results obtained by the DRPA when predicting the presence or absence of DR were accuracy 0.876, sensitivity 84%, specificity 88.5%, positive predictive value 63.8%, and negative predictive value 95.8%. The DRPA Software Retiprogram 4.0 obtained the registration code SAFE CREATIVE 2007144741712 on 14 July 2020 at 6:48 UTC

Testing

The DRPA was tested using the electronic medical records of 40,129 patients with type 2 diabetes extracted from SIDIAP. For this study, data from 11 years of follow-up were extracted. First, the DRPA predicted the presence or absence of DR for each diabetic patient based on the 9 clinical variables needed to feed the algorithm. Then, the same four blinded retina specialists provided a definitive result by interpreting their fundus examinations. Finally, we calculated the agreement between ophthalmologists and the DRPA in predicting the presence of DR.

4.3. Statistical Methods

Data were analyzed using the SPSS statistical software package (software IBM® SPSS® version 25.0, IBM Corp., Armonk, NY, USA). For qualitative data, frequency and percentage analysis were used in each category and were compared using the χ^2 test. Quantitative variables are shown as mean ± SD when normally distributed, otherwise the median and interquartile range are shown. In normal distributions, quantitative variables were compared using parametric tests; the Student's *t*-test was used when it was intended to compare the mean of 2 groups, or the ANOVA analysis was used if there were more than two groups. The receiver operating characteristic (ROC) analysis objectified the variables that best fitted the model prediction. Finally, the variables that were dependent on diabetes duration were assessed through the Cox survival analysis.

We measured the performances of both the AIRS and DRPA using a confusion matrix in which the categories assigned by the algorithms were compared to those given by the retina experts which was considered the ground truth. There were four possible combinations: true positives (TPs) or correct positive assignments, true negatives (TNs) or correct negative assignments, false positives (FPs) or incorrect positive assignments, and false negatives (FNs) or incorrect negative assignments. The performance of the algorithms was evaluated through the following scores: accuracy (ACC), sensitivity (S), specificity (SP), positive predictive value (PPV), negative predictive value (NPV), positive false discovery rate or type 1 error (α), negative false discovery rate or type 2 error (β), and the area under the curve (AUC). *p* < 0.05 was considered statistically significant.

5. Results

We show the results obtained by the AIRS when reading and classifying retinal images compared to those provided by retina experts, both in our database and in Messidor-2. Next, we present the results obtained by the DRPA when predicting the presence or absence of DR compared to the categories assigned by retina specialists.

5.1. Testing the AIRS in Our Database

The AIRS read 15,297 retinal pictures and only discarded 55 (0.91%) due to poor image quality. Our database contained 12,538 images with no DR (81.9%), 938 with mild DR (6.1%), 1109 with moderate DR (7.2%), and 722 with sight-threatening DR (4.7%).

The algorithm correctly classified 99.9%, 95%, 96%, and 88.7% of those retinographies with no, mild, moderate, and severe DR, respectively. The performance of the AIRS is shown in Table 6. We show an example of the ability of AIRS to classify a retina picture as severe (Figure 2).

Table 6. Confusion matrix. Performance of the artificial-intelligence-based reading system (AIRS) in classifying retinographies into each levels of DR severity compared to the classification given by retina experts in our sample of diabetics.

		Predictions Given by the AIRS			
		0	1	2	3
Classification provided by Ophthalmologists	0	12.71	12	0	0
	1	41	892	5	0
	2	9	35	1.065	0
	3	0	17	64	641
Total		12.621	956	1.129	641

0: No DR; 1: mild DR; 2: moderate DR; 3: severe and proliferative DR. The algorithm correctly classified 91% of retinographies. It only misclassified 12 (0.01%), 46 (4.9%), 44 (3.9%), and 81 (11.2%) of those retinographies with no, mild, moderate, and severe DR, respectively.

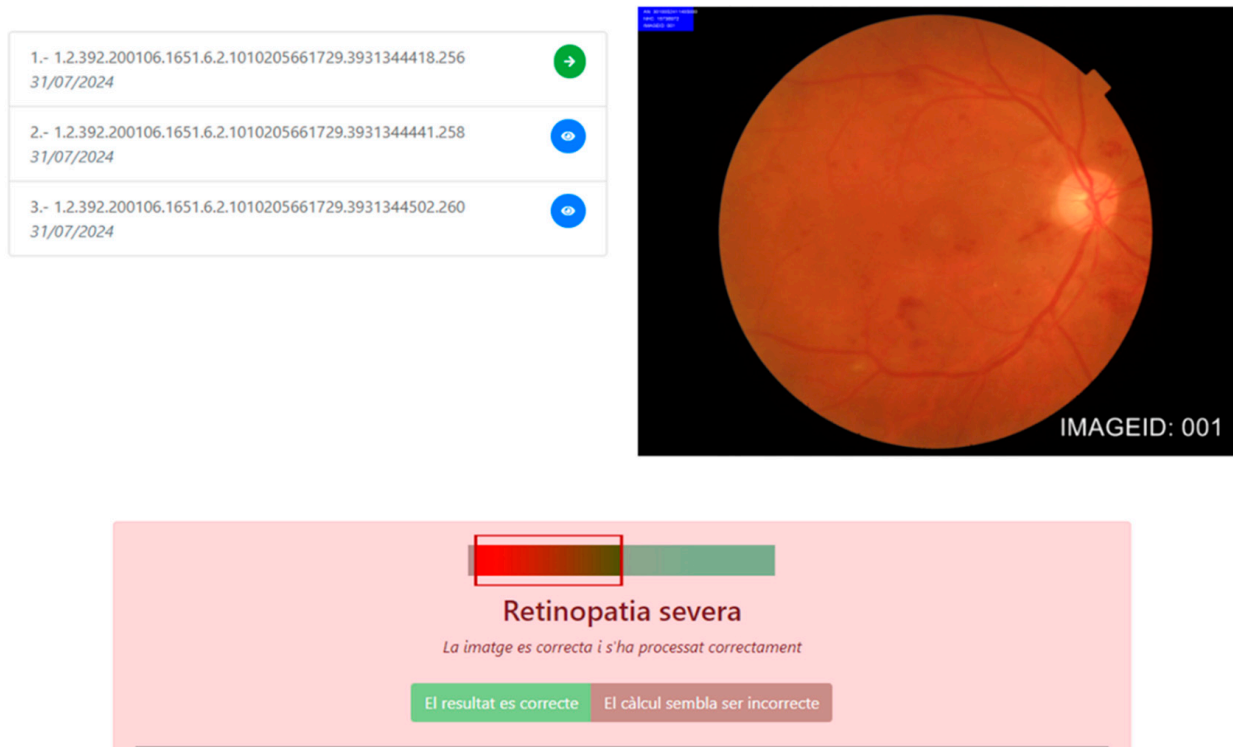


Figure 2. The artificial intelligence-based reading system interface. The algorithm classified this example into severe diabetic retinopathy (Retinopatia severa in Catalan language) due to the presence of multiple retinal hemorrhages.

Table 7 shows the ability of the AIRS to differentiate those retinographies with non-referable DR (no DR + mild DR) into those with referable DR (moderate or severe DR). The

performance of the algorithm in differentiating those retinographies without RDR from those with RDR was ACC = 98.66, S = 96.7, SP = 99.82, PPV = 99.01, NPV = 98.01, and AUC = 0.958.

Table 7. Performance of the reading algorithm when differentiating images with nonreferable DR from those with referable DR in our diabetic database.

		Predictions Given by the AIRS	
		Nonreferable DR	Referable DR
Classification provided by ophthalmologists	Nonreferable DR	13,463	64
	Referable DR	58	1706
		13,521	1770

The AIRS correctly identified 13,463 (99.5%) and 1706 (96.8%) retinographies with nonreferable and referable DR, respectively. In 64 (3.2%) cases, the AIRS identified RDR when it was not present and misclassified 58 (3.7%) cases as not having RDR when it was present.

5.2. Testing the AIRS with Messidor-2

The gradeability test performed by the AIRS on the 1200 retinographies from Messidor-2 found all of them gradable (100%). This was due to the high quality of all the retinal images in that dataset. The AIRS correctly classified 97.6%, 63.3%, 89.3%, and 95.8% of those retinographies contained in the Messidor-2 database with no, mild, moderate, and severe DR, respectively. The algorithm only misclassified 3.4% of these fundus images as having a milder degree of DR than they did (Table 8).

Table 8. Confusion matrix. Performance of the reading algorithm regarding its ability to classify the retinographies extracted from Messidor-2 into levels of severity compared to the classification provided by Messidor-2.

		Predictions Given by the AIRS			
		0	1	2	3
Classification provided by MESSIDOR-2	0	610	35	0	0
	1	13	143	7	0
	2	2	15	116	10
	3	0	4	7	238
Total		625	197	130	248

0: No DR; 1: mild DR; 2: moderate DR; 3: severe and proliferative DR.

We studied the ability of the AIRS in differentiating those retinographies with nonreferable DR (no DR or mild DR) from those with referable DR (moderate and severe DR). The AIRS identified 373 retinographies that presented referable DR (96.5%) and differentiated them from 809 (99.6%) which had nonreferable DR. The AIRS had only 0.6% and 3.3% of false positives and negatives, respectively (Table 9).

Table 9. Confusion matrix. The ability of the AIRS to analyze retinographies with and without referable DR extracted from the Messidor-2 database.

		Predictions Given by the AIRS	
		Nonreferable DR	Referable DR
Classification provided by MESSIDOR-2	Non-referable DR	809	5
	Referable DR	13	373
		822	378

The algorithm only misclassified 0.16% and 3.3% of nonreferable and referable DR, respectively.

The performance scores when differentiating retina pictures with RDR from those without was S = 94.64%, SP = 99.14%, PPV = 90.54%, NPV = 99.53%, ACC = 96.78%, and AUC = 0.918.

5.3. Testing the Diabetes Retinopathy Prediction Algorithm

5.3.1. Clinical and Demographic Data at Baseline

To test the DRPA we had access to 265,388 eligible type 2 diabetic patients from the SIDIAP database, of whom 40,129 patients had the nine clinical variables necessary to feed the DRPA. A total of 22,859 (58%) male patients were included. The mean age was 68.12 ± 10.39 years and the mean duration of diabetes was 9.21 ± 5.51 . There were 31,019 (77.3%) patients who controlled their diabetes with oral agents, 4691 (12.6%) with insulin, and 4419 (10.1%) with diet. One-third of them (32%) had arterial hypertension. The mean microalbuminuria and eGFR were 257.3 ± 122.83 and 73.09 ± 15.24 , respectively. Table 10 shows the mean values of the clinical and demographic variables of the sample used to test the predictive algorithm for the presence of diabetic retinopathy at baseline.

Table 10. Variables used to test the predictive algorithm for the presence of diabetic retinopathy.

Independent Variables	Data
Gender	
• Men	22,859 (58%)
• Women	17,270 (42%)
Age (years)	68.12 ± 10.39 (33–99)
BMI (kg/m ²)	27.28 ± 5.2 (18–36.81)
Blood pressure control:	
• Normal arterial tension	27,288 (68%)
• Arterial hypertension	12,841 (32%)
Glycosylated hemoglobin (%)	7.76 ± 1.61 (4.8–16.6)
Diabetes duration (years)	9.21 ± 5.51 (0.9–55.98)
Microalbuminuria (mg/24 h)	257.3 ± 122.83 (16.23–3155.72)
Diabetes treatment	
• Diet	4419 (10.1%)
• Oral agents	31,019 (77.3%)
• Insulin	4691 (12.6%)
eGFR	73.09 ± 15.24 (60.05–83.25)

BMI: body mass index; eGFR: estimated glomerular filtration rate. In the sample used to test the predictive algorithm, there were more men than women. They were overweight and had reasonably good metabolic control. Most patients were controlled with oral agents. The mean duration of diabetes was less than 10 years. Three-quarters of patients had good blood pressure control and the mean microalbuminuria was almost in the nephrotic range.

At the beginning of the study, there were 40,129 T2DM patients, of whom 36,758 had no DR (91.6%) and 3371 had mild DR (8.4%). After 11 years of follow-up, 2924 patients (7.2%) developed some type of DR: 4293 patients (11.5%) showed mild DR, 1398 (3.5%) showed moderate DR, 199 (0.5%) showed severe DR, 164 showed proliferative DR (0.29%), and 241 showed diabetic macular edema (1.05%). In total, 6295 patients (15.6%) presented with some type of DR at the end of the study (Table 11).

Table 11. Number and proportion of diabetic patients with any type of DR at baseline and after a 11-year period of follow up.

	Type of DR at Baseline		Type of DR at the End of the Study	
0	36,758	(91.6%)	33,898	(85.5%)
1	3371	(8.4%)	4293	(11.7%)
2			1398	(3.5%)
3			199	(0.5%)
4			164	(0.29%)
5			241	(1.05%)
Patients with any DR			6295	(15.6%)

0: No DR; 1: mild DR; 2: moderate DR; 3: severe; 4: proliferative DR; 5: diabetic macular edema. During the 11-year period of follow-up, 3.3%, 3.5%, 0.5%, and 0.29% of patients developed mild, moderate, severe, and proliferative DR, respectively. In addition, 1% developed diabetic macular edema.

5.3.2. Performance of the Predictive Diabetic Retinopathy Algorithm

The algorithm correctly predicted 36,196 and 2795 patients with no and any DR, respectively. In contrast, it predicted that 774 would develop DR when they did not have it (false positives). Finally, it considered that 1090 patients would not develop DR when they did have it (false negatives). $S = 0.89$, $SP = 0.98$, $PPV = 0.79$, $NPV = 0.98$, and $ACC = 0.96$. The AUC was 0.92 for any type of DR.

5.4. Patient Journey for the Early Detection of Diabetic Retinopathy in Primary Care

Primary care teams are the health professionals who most frequently care for diabetic patients. This fact should be used to involve primary care teams in eye screening at any opportunistic visit.

Following this scheme, the patient journey is as follows: T2DM patients visit their primary care teams to have their retinographies taken and estimate their risk of developing DR by entering the nine clinical variables to feed DRPA. Then, the clinical data are uploaded to the cloud where they are analyzed by both the AIRS and the DRPA, thus allowing primary care teams to know in real time whether their patients have DR or not and to know their risk of developing it. Only those cases with either DR or high risk of developing it would be referred to the ophthalmologist. The others would be reevaluated according to their risk at different time intervals in the primary care level (Figure 3).

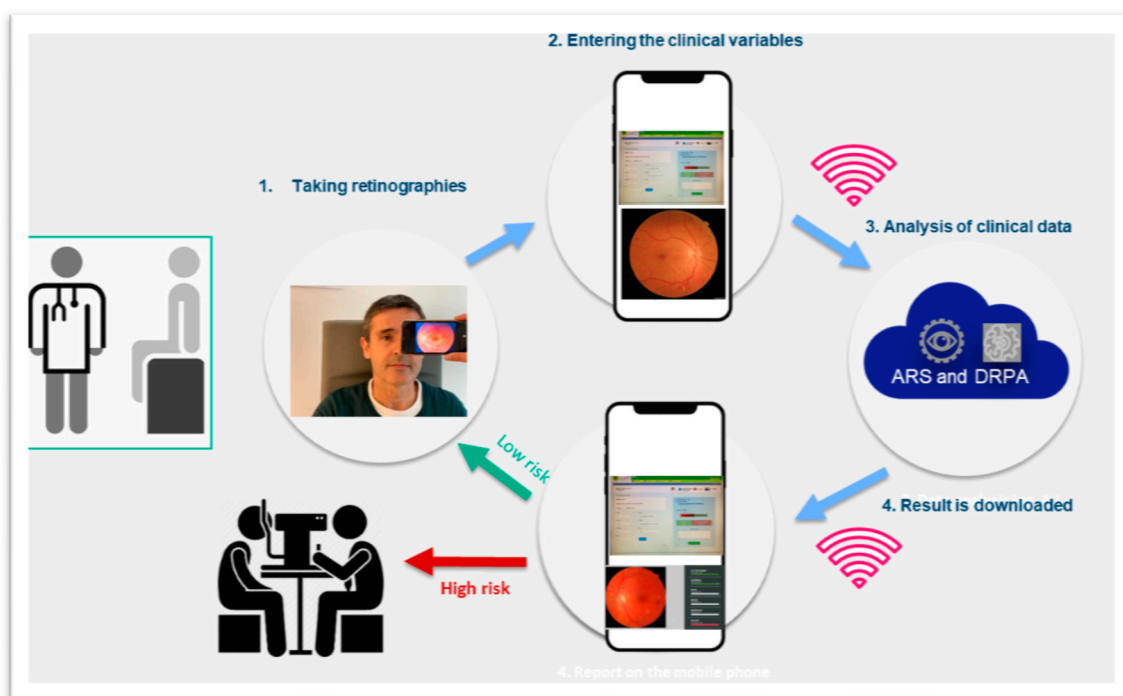


Figure 3. Patient journey according to our eye screening system for the early detection of diabetic retinopathy in primary care. Primary care teams take the retinographies and enter the 9 clinical variables to feed the predictive algorithm. Then, all the data are uploaded to the cloud where the reading and predictive algorithms (the AIRS and the DRPA) evaluate the risk of each patient of developing DR. Only those patients at high risk would be referred to the ophthalmologist. The others would be reevaluated at the primary care level.

6. Discussion

This study was aimed at testing two complementary algorithms for the early detection of DR in T2DM patients. On the one hand, the ability of the artificial-intelligence-based reading system (AIRS) to read and classify the retinographies of patients with DM2 into severity levels using the four standard Messidor-2 categories was tested. On the other hand, the ability of the diabetic retinopathy prediction algorithm (DRPA) to predict the presence or absence of any type of DR based on nine clinical variables was tested. This system was devised to ease the burden of screening for DR and to be applied mainly in primary care.

To build and train the AIRS, 88,692 EyePACS [24] retinal images were used along with a sample of 15,123 labeled retinal images from our database. Then, it was validated using 5000 retinographies from our database and, finally, it was tested using 15,297 retinographies from our database and 1200 retinal images extracted from Messidor-2.

Our AIRS obtained a sensitivity of 96.7%, a specificity of 99%, a positive predictive value of 99%, a negative predictive value of 99.7%, and an area under the curve of 0.958 when differentiating those photographs without DR or mild DR (nonreferable DR) from those with moderate or severe DR (referable DR) in our diabetic population. When analyzing the Messidor-2 database, the AIRS yielded a sensitivity of 94.6%, a specificity of 99%, a positive predictive value of 90.5%, a negative predictive value of 99.5%, and an area under the curve of 0.918. The false positives shown by the AIRS were 0.08% and 0.86% and the false negatives were 0.1% and 3.3% in our diabetic population and Messidor-2, respectively. Therefore, the AIRS was less accurate when analyzing fundus images extracted from Messidor-2 compared to those in our database. It is noteworthy that our reading algorithm was trained from fundus images taken with a single nonmydriatic fundus camera (TOPCON TRC-NW6S), whereas the fundus images contained in the Messidor-2 database were obtained with several fundus camera models. The difference in pixel definition between

the fundus images could explain the better performance of AIRS in our diabetic population compared to Messidor-2.

In 2018, the IDX-DR[®] reading algorithm was the first to obtain FDA marketing approval for the automatic detection of DR, achieving a sensitivity of 96.8% and a specificity of 59.4% in detecting referable DR. Since then, several groups have built their algorithms, including EyeArt[®] and Retmarker[®], which have obtained CE Class IIa marking. EyeArt[®] showed a sensitivity of 91.3%, a specificity of 91.1%, and an AUC of 0.96 in classifying those retinographs from the UK National Health System with referable DR [17]. Retmarker[®] also used retinographies from the NHS for its validation, achieving a sensitivity of 73% and a specificity of 85% in detecting any type of DR [35]. It is fair to conclude that our AIRS performed well in both samples in detecting RDR, compared to those that have already been approved by the FDA or that have obtained the CE Class IIa mark. Identifying patients with this level of retinopathy is crucial for referring them to a retina specialist as early as possible and, thus, preventing vision loss and blindness caused by this condition.

The DRPA was tested by conducting a retrospective study of 11 years in a population of 40,129 T2DM patients. It obtained an AUC of 0.92, a sensitivity of 89%, and, more importantly, a specificity of 98% and a negative predictive value of 98%, meaning that the algorithm predicts with sufficient certainty those patients who, due to their clinical characteristics, are not going to develop DR. It allows us to identify patients at low risk of developing DR who would be safe being scanned less frequently, thus freeing up resources for targeting the more urgent cases. When comparing our results with previous studies, it is important to note that only three authors have published work on the development of prediction algorithms: Aspelund [24], Scanlon [27], and Broadbent [28].

The model developed by Aspelund was constructed using the EHR from the Diabetic Retinopathy database at the Ophthalmology Department of Aarhus University Hospital (Denmark) over 20 years. They used five clinical variables: the current age, diabetes duration, systolic blood pressure, HbA1c%, and type of diabetes. After feeding the algorithm with the five variables, it recommends a screening time interval ranging from 6 to 60 months for each diabetic patient, depending on their risk of developing DR. The algorithm was tested in Spain [25] and obtained an AUC of 0.74. In the Netherlands, the algorithm was tested by Van der Heijden [36], who obtained an AUC of 0.83, and in the UK, Lund [26] reported an AUC of 0.83 when predicting the onset of DR for T2DM patients.

The model developed by Scanlon was validated using the EHR of T2DM patients from Gloucestershire, UK. This algorithm was constructed from only three independent variables: age, total cholesterol, and HbA1c%. It obtained an AUC of 0.77 for predicting the development of sight-threatening diabetic retinopathy.

Finally, the model developed by Broadbent was built using seven clinical variables: age at diagnosis, gender, diabetes duration, systolic blood pressure, total cholesterol, HbA1c%, and DR status, to predict the risk of developing sight-threatening diabetic retinopathy. It was tested in the UK and demonstrated a sensitivity, a specificity, and an AUC of 61%, 93%, and 0.88, respectively. It is important to note that all these three algorithms predict only the development of sight-threatening diabetic retinopathy, while our model predicts the presence of any type of DR.

The strengths of the present study are the good performance of our AIRS, not only in our fundus imaging database but also in the public Messidor-2 database. Furthermore, our DRPA not only predicted RDR but any type of DR, and was a very reliable tool to rule out the presence of DR. Limitations of the present study are that we tested our AIRS using fundus images from predominantly Caucasian patients. Therefore, it should be tested on larger samples of patients with greater ethnic diversity. It is worth mentioning that the performance of our AIRS was not as accurate when reading Messidor-2 fundus images compared to those of our population. This is a clear limitation for exporting our algorithm directly to other populations. This fact is probably due to the different image resolutions provided by the different fundus cameras and, therefore, our AIRS must be previously trained and validated in each population that it is intended to be applied on. Another

aspect to consider is the different fundus characteristics that different populations far from each other may have, which would force us to validate our algorithm again if we wanted to apply it. Furthermore, the DRPA algorithm was tested using medical records that were collected some years earlier. This fact led us to choose only those medical records from the Catalan Health System Diabetic Patient Database (SIDIAP) in which the data collection process follows a strict protocol and helps to build trust in them.

7. Conclusions

The AIRS can be used as a reliable tool for identifying those T2DM patients with referral DR. It allows us to reduce the time needed to read the images and, thus, screen many more patients annually for DR. As a result, those patients with referral DR would be treated early. The DRPA identified those patients who are not going to develop DR, which allows us to extend the intervals of screening for DR with greater security. If both algorithms were applied in primary care, nurses and family physicians would quickly be able to refer those patients at risk of vision loss and, thus, improve their visual prognosis.

Author Contributions: M.B.-B. contributed to study conception and design, reviewed the statistical analysis, wrote the discussion, and edited the manuscript, contributing to the final approval of the version sent for publication. B.F.-P. contributed to collecting research data from SIDIAP, wrote the discussion, contributing to the final approval of the version sent for publication. P.R.-A. contributed to study conception and design, reviewed the statistical analysis, wrote the discussion, and edited the manuscript, contributing to the final approval of the version sent for publication. R.N.-G. contributed to ophthalmological data collection, diagnosed diabetic retinopathy and nephropathy, and carried out the laboratory procedures, contributing to the final approval of the version sent for publication. A.H.-C. contributed to collecting data and to the final approval of the version sent for publication. A.B.-P. contributed to collecting data, contributing to the final approval of the version sent for publication. M.L.-S. contributed to retinal picture acquisition and to the final approval of the version sent for publication. C.M.-L. contributed to retinal picture acquisition and to the final approval of the version sent for publication. M.V. contributed to study design, contributing to the final approval of the version sent for publication. A.M.-P. contributed to study design and the statistical analysis, interpreted research data, and contributed to the interpretation of the study findings, contributing to the final approval of the version sent for publication. All authors have read and agreed to the published version of the manuscript.

Funding: EIT health. Grant number: 230 123. This project has received cofunding from the European Institute of Innovation and Technology (EIT) Health and the European Union under Grant Agreement 230123.

Institutional Review Board Statement: CEIM IISPV (Institut d'Investigació Sanitària Pere Virgili) approved the present study (approval code RetinaRead-Risk, protocol version 1. 3 October 2022, Reference number CEIM: 071/2022) which was carried out in accordance with the revised guidelines of the Declaration of Helsinki.

Informed Consent Statement: All patients included in the study were previously informed about its objective. Once their commitment to participate in the follow-up of the study was obtained, they were asked to sign the informed consent. They will be informed about the results obtained at the end of the study.

Data Availability Statement: The database used and analyzed is available from the corresponding author on research request.

Acknowledgments: We thank all healthcare personnel involved in the screening for the early detection of DR. We thank general endocrinologists, opticians, nurses, and family doctors in our area who help us to continue performing the eye screening which has been ongoing since 2007. We thank all the patients who decided to participate in the present study. We also thank Phil Hoddy for his language assistance and for editing and correcting the English text.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ogurtsova, K.; Da Rocha Fernandes, J.D.; Huang, Y.; Linnenkamp, U.; Guariguata, L.; Cho, N.H.; Cavan, D.; Shaw, J.E.; Makaroff, L.E. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res. Clin. Pract.* **2017**, *128*, 40–50. [\[CrossRef\]](#)
2. Soriguer, F.; Goday, A.; Bosch-Comas, A.; Bordiú, E.; Calle-Pascual, A.; Carmena, R.; Casamitjana, R.; Castaño, L.; Castell, C.; Catalá, M.; et al. Prevalence of diabetes mellitus and impaired glucose regulation in Spain: The Di@bet.es Study. *Diabetologia* **2012**, *55*, 88–93. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Teo, Z.L.; Tham, Y.-C.; Yu, M.; Chee, M.L.; Rim, T.H.; Cheung, N.; Bikbov, M.M.; Wang, Y.X.; Tang, Y.; Lu, Y.; et al. Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045: Systematic Review and Meta-analysis. *Ophthalmology* **2021**, *128*, 1580–1591. [\[CrossRef\]](#)
4. Lee, L.J.; Yu, A.P.; Cahill, K.E.; Oglesby, A.K.; Tang, J.; Qiu, Y.; Birnbaum, H.G. Direct and indirect costs among employees with diabetic retinopathy in the United States. *Curr. Med Res. Opin.* **2008**, *24*, 1549–1559. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Happich, M.; Reitberger, U.; Breitscheldel, L.; Ulbig, M.; Watkins, J. The economic burden of diabetic retinopathy in Germany in 2002. *Graefes Arch. Clin. Exp. Ophthalmol.* **2008**, *246*, 151–159. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Ghanchi, F.; the Diabetic Retinopathy Guidelines Working Group. The Royal College of Ophthalmologists' clinical guidelines for diabetic retinopathy: A summary. *Eye* **2013**, *27*, 285–287. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Wu, C.M.; Wu, A.M.; Young, B.K.; Wu, D.J.; Margo, C.E.; Greenberg, P.B. An Appraisal of Clinical Practice Guidelines for Diabetic Retinopathy. *Am. J. Med Qual.* **2016**, *31*, 370–375. [\[CrossRef\]](#)
8. López, M.I.; Díez, A.; Vellilla, S.; Rueda, A.; Alvarez, A.; Pastor, J.C. Prevalence of diabetic retinopathy and eye care in a rural area of Spain. *Ophthalmic Epidemiol.* **2002**, *9*, 205–214. [\[CrossRef\]](#)
9. Wong, T.Y.; Bressler, N.M. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA* **2016**, *316*, 2366–2367. [\[CrossRef\]](#)
10. Ting, D.S.W.; Cheung, C.Y.-L.; Lim, G.; Tan, G.S.W.; Quang, N.D.; Gan, A.; Hamzah, H.; Garcia-Franco, R.; Yeo, I.Y.S.; Lee, S.Y.; 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*, 2211–2223. [\[CrossRef\]](#)
11. Abramoff, M.D.; Lavin, P.T.; Birch, M.; Shah, N.; Folk, J.C. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit. Med.* **2018**, *1*, 39. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Scotland, G.S.; McNamee, P.; Philip, S.; Fleming, A.D.; Goatman, K.A.; Prescott, G.J.; Fonseca, S.; Sharp, P.F.; Olson, J.A. Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in Scotland. *Br. J. Ophthalmol.* **2007**, *91*, 1518–1523. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Tufail, A.; Rudisill, C.; Egan, C.; Kapetanakis, V.V.; Salas-Vega, S.; Owen, C.G.; Lee, A.; Louw, V.; Anderson, J.; Liew, G.; et al. Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders. *Ophthalmology* **2017**, *124*, 343–351. [\[CrossRef\]](#) [\[PubMed\]](#)
14. O'keeffe, D.; Riordan, F.; Harkins, V.; Kearney, P.; Mc Hugh, S. Predictors of attendance at diabetic retinopathy screening among people with type 2 diabetes: Secondary analysis of data from primary care. *Prim. Care Diabetes* **2021**, *15*, 1086–1094. [\[CrossRef\]](#)
15. Van Der Heijden, A.A.; Abramoff, M.D.; Verbraak, F.; Van Hecke, M.V.; Liem, A.; Nijpels, G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. *Acta Ophthalmol.* **2018**, *96*, 63–68. [\[CrossRef\]](#)
16. Schmidt-Erfurth, U.; Sadeghipour, A.; Gerendas, B.S.; Waldstein, S.M.; Bogunović, H. Artificial intelligence in retina. *Prog. Retin. Eye Res.* **2018**, *67*, 1–29. [\[CrossRef\]](#)
17. Bhaskaranand, M.; Ramachandra, C.; Bhat, S.; Cuadros, J.; Nittala, M.G.; Sadda, S.R.; Solanki, K. The value of automated diabetic retinopathy screening with the EyeArt system: A study of more than 100,000 consecutive encounters from people with diabetes. *Diabetes Technol. Ther.* **2019**, *21*, 635–643. [\[CrossRef\]](#)
18. Ribeiro, M.L.; Nunes, S.G.; Cunha-Vaz, J.G. Microaneurysm turnover at the macula predicts risk of development of clinically significant macular edema in persons with mild nonproliferative diabetic retinopathy. *Diabetes Care* **2013**, *36*, 1254–1259. [\[CrossRef\]](#)
19. González-Gonzalo, C.; Sánchez-Gutiérrez, V.; Hernández-Martínez, P.; Contreras, I.; Lechanteur, Y.T.; Domanian, A.; van Ginneken, B.; Sánchez, C.I. Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration. *Acta Ophthalmol.* **2020**, *98*, 368–377. [\[CrossRef\]](#)
20. Wong, D.C.S.; Kiew, G.; Jeon, S.; Ting, D.; Wong, D.; Kiew, G. Singapore Eye Lesions Analyzer (SELENA): The Deep Learning System for Retinal Diseases. In *Artificial Intelligence in Ophthalmology*, 1st ed.; Grzybowski, A., Ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 177–185.
21. Li, Z.; Keel, S.; Liu, C.; He, Y.; Meng, W.; Scheetz, J.; Lee, P.Y.; Shaw, J.; Ting, D.; Wong, T.Y.; 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*, 2509–2516. [\[CrossRef\]](#)
22. Ruamviboonsuk, P.; Krause, J.; Chotcomwongse, P.; Sayres, R.; Raman, R.; Widner, K.; Campana, B.J.L.; Phene, S.; Hemarat, K.; Tadarati, M.; et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. *NPJ Digit. Med.* **2019**, *2*, 25. [\[CrossRef\]](#)
23. Soliz, P. Multicenter, head-to-head, real-world validation study of seven automated artificial intelligence diabetic retinopathy screening systems. *Diabetes Care* **2021**, *44*, 1168–1175. [\[CrossRef\]](#)

24. Aspelund, T.; Þórisdóttir, Ó.; Ólafsdóttir, E.; Guðmundsdóttir, A.; Einarisdóttir, A.B.; Mehlsen, J.; Einarsson, S.; Pálsson, Ó.; Einarsson, G.; Bek, T.; et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. *Diabetologia* **2011**, *54*, 2525–2532. [[CrossRef](#)]
25. Soto-Pedre, E.; Pinies, J.A.; Hernaez-Ortega, M.C. External validation of a risk assessment model to adjust the frequency of eye-screening visits in patients with diabetes mellitus. *J. Diabetes Its Complicat.* **2015**, *29*, 508–511. [[CrossRef](#)] [[PubMed](#)]
26. Lund, S.H.; Aspelund, T.; Kirby, P.; Russell, G.; Einarsson, S.; Pálsson, O.; Stefánsson, E. Individualised risk assessment for diabetic retinopathy and optimisation of screening intervals: A scientific approach to reducing healthcare costs. *Br. J. Ophthalmol.* **2016**, *100*, 683–687. [[CrossRef](#)] [[PubMed](#)]
27. Scanlon, P.H.; Aldington, S.J.; Leal, J.; Luengo-Fernandez, R.; Oke, J.; Sivaprasad, S.; Gazis, A.; Stratton, I.M. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. *Heal. Technol. Assess.* **2015**, *19*, 1–116. [[CrossRef](#)]
28. Broadbent, D.M.; Sampson, C.J.; Wang, A.; Howard, L.; Williams, A.E.; Howlin, S.U.; Appelbe, D.; Moitt, T.; Cheyne, C.P.; Rahni, M.M.; et al. Individualised screening for diabetic retinopathy: The ISDR study—rationale, design and methodology for a randomised controlled trial comparing annual and individualised risk-based variable-interval screening. *BMJ Open* **2019**, *9*, e025788. [[CrossRef](#)]
29. Decencière, E.; Zhang, X.; Cazuguel, G.; Lay, B.; Cochener, B.; Trone, C.; Gain, P.; Ordonez, R.; Massin, P.; Erginay, A.; et al. feedback on a publicly distributed image database: The Messidor database. *Image Anal. Ster.* **2014**, *33*, 231. [[CrossRef](#)]
30. Cuadros, J.; Bresnick, G. EyePACS: An adaptable telemedicine system for diabetic retinopathy screening. *J. Diabetes Sci. Technol.* **2009**, *3*, 509–516. [[CrossRef](#)]
31. de la Torre, J.; Puig, D.; Valls, A. Weighted kappa loss function for multi-class classification of ordinal data in deep learning. *Pattern Recognit. Lett.* **2018**, *105*, 144–154. [[CrossRef](#)]
32. Romero-Aroca, P.; Verges-Puig, R.; de la Torre, J.; Valls, A.; Relaño-Barambio, N.; Puig, D.; Baget-Bernaldiz, M. Validation of a Deep Learning Algorithm for Diabetic Retinopathy. *Telemed. e-Health* **2020**, *26*, 1001–1009. [[CrossRef](#)] [[PubMed](#)]
33. Romero-Aroca, P.; Valls, A.; Moreno, A.; Sagarra-Alamo, R.; Basora-Gallisa, J.; Saleh, E.; Baget-Bernaldiz, M.; Puig, D. A Clinical Decision Support System for Diabetic Retinopathy Screening: Creating a Clinical Support Application. *Telemed. e-Health* **2019**, *25*, 31–40. [[CrossRef](#)] [[PubMed](#)]
34. Romero-Aroca, P.; Verges-Pujol, R.; Santos-Blanco, E.; Maarof, N.; Valls, A.; Mundet, X.; Moreno, A.; Galindo, L.; Baget-Bernaldiz, M. Validation of a diagnostic support system for diabetic retinopathy based on clinical parameters. *Transl. Vis. Sci. Technol.* **2021**, *10*, 17. [[CrossRef](#)]
35. Ribeiro, L.; Oliveira, C.M.; Neves, C.; Ramos, J.D.; Ferreira, H.; Cunha-Vaz, J. Screening for diabetic retinopathy in the central region of portugal. added value of automated ‘disease/no disease’ grading. *Ophthalmologica* **2015**, *233*, 96–103. [[CrossRef](#)] [[PubMed](#)]
36. van der Heijden, A.A.W.A.; Walraven, I.; Riet, E.V.; Aspelund, T.; Lund, S.H.; Elders, P.; Polak, B.C.P.; Moll, A.C.; Keunen, J.E.E.; Dekker, J.M.; et al. Validation of a model to estimate personalised screening frequency to monitor diabetic retinopathy. *Diabetologia* **2014**, *57*, 1332–1338. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.