ARTIFICIAL INTELLIGENCE IN DIABETIC RETINOPATHY SCREENING. A REVIEW

Straňák Z.*, Penčák M.*, Veith M.

Department of Ophthalmology, 3rd Faculty of Medicine, Charles University and University Hospital Kralovske Vinohrady, Prague, Czech Republic

***The first and second authors contributed equally to this article.**

The authors of the study declare that no conflict of interest exists in the compilation, theme and subsequent publication of this professional communication, and that it is not supported by any pharmaceuticals company. The study has not been submitted to any other journal or printed elsewhere, with the exception of congress abstracts and recommended procedures.

Presented at the XVIII Congress of the Czech Vitreoretinal Society in Brno 2018 and XXX. Ophthalmology Day in Ústí nad Labem 2019

Received: 19 October 2020
Accepted: 8 December 2020
Available online: 10 March 2021

MUDr. Zbyněk Straňák
Oftalmologická klinika FNKV
Šrobárova 50
100 34 Praha 10
E-mail: zbynek.stranak@fnkv.cz

SUMMARY

Objective: The aim of this comprehensive paper is to acquaint the readers with evaluation of the retinal images using the artificial intelligence (AI). Main focus of the paper is diabetic retinophaty (DR) screening. The basic principles of the artificial intelligence and algorithms that are already used in clinical practice or are shortly before approval will be described.
Methodology: Describing the basic characteristics and mechanisms of different approaches to the use of AI and subsequently literary minireview clarifying the current state of knowledge in the area.
Results: Modern systems for screening diabetic retinopathy using deep neural networks achieve a sensitivity and specificity of over 80 % in most published studies. The results of specific studies vary depending on the definition of the gold standard, number of images tested and on the evaluated parameters.
Conclusion: Evaluation of images using AI will speed up and streamline the diagnosis of DR. The use of AI will allow to keep the quality of the eye care at least on the same level despite the raising number of the patients with diabetes.
Key words: artificial intelligence, screening, diabetic retinopathy

Čes. a slov. Oftal., 77, 2021, No.5, p. 223–230

INTRODUCTION

In recent years, the use of artificial intelligence has found itself in the limelight in all fields of human medicine, from radiology to ophthalmology [1]. One of the examples in which an intelligent program has surpassed a human colleague is the algorithm for detecting skin cancer, which is able to determine this clinical unit more precisely than an experienced dermatologist [2]. Furthermore, artificial intelligence (AI) here is faster and more effective, and by contrast with several years of study and clinical preparation requires only training datasets.

Within the context of image analysis, the retina has a similarly suitable position. Since 1926 we have been able to obtain photographs of the ocular fundus on a commercially available device (the Carl Zeiss Company has precedence) [3]. With improvements in technology, today various types of wide-angle cameras are available, as well as for example hand-held cameras which do not require mydriasis or a regularly trained ophthalmologist, and can therefore be a component of the regular equipment of an outpatient diabetologist, for example [4]. Nevertheless, an image obtained in this manner must still be sent for assessment outside of the clinic of the relevant diabetologist. It is precisely the use of fundus photography within the framework of screening for diabetic retinopathy (DR) that has become the focus of considerable attention recently.

DR is a serious complication of diabetes. It represents a leading cause of blindness among the adult population of productive age [5,6]. According to data published by the Institute of Health Information and Statistics in 2017, there were 930 000 diabetic patients in the Czech Republic [7]. DR was demonstrated in 10.3% of these patients, of whom 24.9% suffered from severe proliferative form. In 0.24% of all diabetics, visual acuity was on the level of blindness.

In some countries, the introduction of nationwide program-

mes for DR screening based on the principles of telemedicine and evaluation of the degree of diabetic retinopathy from photographs of the ocular fundus has led to a reduction of the incidence of blindness caused by diabetic retinopathy [5]. These programmes enable regular and effective screening of a large number of patients with diabetes. They are usually based on a multi-stage assessment of fundus photographs by a group of specially trained evaluators, who are usually not doctors. Nevertheless, the introduction of such programmes is highly demanding, both financially and in terms of personnel [8,9]. A larger expansion could be facilitated by the introduction of computer algorithms in order to assess photographs of the ocular fundus [10,11].

A substantial advance has been achieved in this area recently due to the advances in the development of artificial intelligence programs based on the principles of deep learning and deep neural networks. The technique of deep learning enables advanced processing and assessment of visual and audio information [12]. The most recent published results show that the introduction of these programs into clinical practice could lead to a reduction of the number of necessary evaluators, while preserving the required sensitivity and specificity [10]. The aim of this article is to present an overview of the current possibilities of programs based on the principle of deep learning and deep neural networks in the diagnosis of diabetic retinopathy from photographs of the ocular fundus.

OBJECTIVE

To acquaint readers with an evaluation of retinal images using artificial intelligence, and to present the functioning and results of the systems most widely used today.

METHODS

Research of the literature, focusing on assessment of images within the framework of DR screening with the aid of AI. We evaluated articles published up to June 2020 with the use of the following key words: "diabetic retinopathy screening", "deep learning", "artificial intelligence" and "automated diabetic retinopathy system". We focused primarily on data concerning specificity and sensitivity.

Specificity expresses the capacity of the test to correctly identify negative findings. This is the number of negative results in proportion to negatives plus false positives. If specificity is 100 %, this means that all the negative findings were genuinely negative. If specificity is 67 %, this means that of 3 genuinely negative eyes (e.g. for DR), only 2 eyes were determined as negative.

Sensitivity expresses the capacity of the test to correctly identify positive findings. This is the number of positive results in proportion to positives plus false negatives. If sensitivity is 100%, then no positive finding has escaped identification (e.g. the test indicates 3 eyes as positive, though in reality only 2 are positive). If sensitivity is 67%, this means that of 3 positive eyes, the test determined only 2 as positive.

RESULTS

IDx-DR

IDx-DR (Digital Diagnostics, Coralville, USA) is a hybrid system which uses a number of deep neural networks, of which each is trained for the detection of a certain sign of DR – haemorrhages, microaneurysms, hard exudates etc. [13]. It also contains an algorithm evaluating the quality of the assessed fundus photographs. It is designated for use with the non-mydriatic fundus camera Topcon NW400 (Topcon Corporation, Tokyo, Japan). For evaluation it uses two 45° fundus photographs of each eye – one centred on the fovea and the other on the optic nerve.

The photographs are then sent via the internet to a remote server, where assessment takes place. The program IDx-DR first of all assesses the quality of the photographs, and in the case of sufficient quality then performs an evaluation and sends the result back to the attending doctor. The system evaluates the presence of greater than mild diabetic retinopathy, sight-threatening DR and diabetic macular edema.

A pilot study recruited 900 patients, of whom 819 completed the study and could be fully evaluated. IDx-DR was tested against human evaluators as the gold standard. IDx-DR performed an evaluation from two images of the fundus for each eye. In 76.4 % of cases, sufficient quality of the photographs was attained without the necessity of pharmacological mydriasis. The human evaluators had four stereoscopic images from a wide-angle fundus camera available, covering a surface of standard 7 ETDRS fields, as well as an image from optical coherence tomography (OCT) of the macula. The images were assessed by 3 experienced evaluators. The factors assessed were the presence of greater than mild DR, diabetic macular edema and sight-threatening DR [13]. In this study the program attained a rate of 87.2% sensitivity and 90.7% specificity, and 96.1% of all images could be evaluated. Sensitivity in sight-threatening diabetic retinopathy was as high as 97.4%. The system was also capable of detecting 84% of patients with diabetic macular edema affecting the macula present on OCT only from fundus photographs. The human evaluators succeeded in identifying only 34% of these patients from stereoscopic fundus photographs. On the basis of this study, the possibility of the use of IDx-DR for the detection of DR was approved by the American Food and Drugs Administration (FDA). This represents the first medical device using artificial intelligence to receive this approval [14].

Good results from the pilot study were also confirmed upon the deployment of the program in clinical practice. Upon use in the Netherlands, the system attained a rate of 91.0% sensitivity and 84.0% specificity in a population of 1415 diabetics, although the system determined that only 66.3% of the images could be evaluated. For sight-threatening DR sensitivity was 62.0% and specificity was 95.0% [15].

In another Dutch study, for a population of 1616 patients the system attained 79.4% sensitivity and 93.8% specificity for greater than mild DR [16]. For sight-threatening retinopathy, sensitivity was 100.0% and specificity 97.8%, 17.3% of images could not be evaluated.

In a study conducted in Spain, the system attained

100% sensitivity and 82% specificity for greater than mild DR. For sight-threatening DR, sensitivity was also 100%, with 95% specificity. The system determined that 76% of images were of sufficient quality for evaluation, images of 2680 patients were evaluated [17].

A pilot study in Poland evaluated 450 patients. The IDx-DR system was capable of evaluating 76% of images, with 94% sensitivity and 95% specificity [18].

Selena

Selena (EyRIS, Singapore) was developed by the Singapore Eye Research Institute. In addition to DR, Selena is also capable of identifying patients with age-related macular degeneration and suspected glaucoma [19]. The system is able to evaluate fundus photographs obtained using various types of retinal cameras. In the study, the sensitivity and specificity of the program was comparable with human evaluators – non-doctors who contributed to the evaluation of the images within the framework of a telemedicine program for screening of diabetic retinopathy, which has been introduced in Singapore. The gold standard was evaluation by a retina specialist. The program achieved a rate of 90.5% sensitivity and 91.6% specificity, compared with 88.5% and 99.6% in the case of human evaluators.

The success rate of the program was not influenced by the race, sex or age of the patients. For detection of suspected glaucoma the program had 96.4% sensitivity and 87.2% specificity, for age-related macular degeneration 93.2% sensitivity and 88.7% specificity. The study also compared two scenarios of deployment of the program. In the fully automated scenario, the program itself decided upon the sending of the patient to a doctor. In the semi-automated scenario, patients who had been indicated for treatment by the program were further evaluated by human evaluators. In the fully automated scenario the sum sensitivity for all three diagnoses was 93%, with specificity of 77.5%. In the semi-automated scenario, sensitivity was 91.3% and specificity 99.5%, with the result that 25.3% of the images were sent for evaluation by human evaluators.

Bosch DR Algorithm

The program was created by the firm Robert Bosch GmbH (Gerlingen-Schillerhöhe, Germany). The system distinguishes between three findings – healthy, with presence of DR, and inconclusive. The system uses a portable non-mydriatic Bosch camera. The only published study evaluated 1128 eyes of 564 patients. One non-mydriatic fundus photograph was evaluated for each eye. The gold standard was evaluation of stereophotographs of the ocular fundus in mydriasis in the standard 7 ETDRS fields. The system achieved 91.2% sensitivity and 96.9% specificity. The system was unable to evaluate 3.9% of all photographs. In only 4 patients the system was unable to evaluate a photograph of either eye.

RetmarkerDR

RetmarkerDR (Retmarker, Coimbra, Portugal) has been used in a screening programme in central Portugal since 2011. In the original version this represented a system for detecting signs of DR on the basis of machine learning. The program is capable of comparing photographs of the ocular fundus from different visits, and evaluating changes in the incidence of microaneurysms, which could be a predictive factor for the development of diabetic macular edema [20]. In the original version the system attained 73 % sensitivity for DR, 85% for significant DR and 97.9% for proliferative DR. Specificity for significant DR was 52.3% [10].

In the latest version, new algorithms were implemented into the program, based on techniques of deep learning. This led to a significant improvement in specificity. In the tested set of photographs, the innovated algorithm achieved sensitivity of 97% in comparison with 96.5% in the original version of the program. Specificity improved from 60.2% in the original version to 88% in the new version of the program [21].

Google Inception

Researchers from the Google company developed artificial intelligence architecture based on the principle of deep neural networks with the code name Inception, which is specially designed for image recognition [22]. This architecture has subsequently been improved several times [23]. Different generations of this architecture have been trained by several research teams to recognise signs of DR. Gulshan et al. tested this system on a publicly available set of fundus photographs of patients with diabetes, as well as on a sample of fundus photographs of patients from a DR screening programme in the USA and India [24]. They tested the system in two configurations – in one for high sensitivity, in the second for high specificity. The gold standard was a majority decision of a group of certified ophthalmologists. The system evaluated the presence of medium and worse DR, severe DR (advanced non-proliferative and proliferative DR), and diabetic macular edema.

Upon an evaluation of the publicly available set of photographs, in configuration for high specificity, the program attained 87.0% sensitivity and 98.5% specificity, and for the sample from the screening programme 90.3% sensitivity and 98.1% specificity. In configuration for high sensitivity, the program attained 96.1% sensitivity and 93.9% specificity for the publicly available set of photographs, and 97.5% and 93.4% respectively for the sample from the screening programme.

This algorithm was further improved [25] and tested in real deployment within the framework of a diabetic retinopathy screening programme in Thailand [26]. The performance of the program was compared with that of the human evaluators who contributed to the screening programme in Thailand. A total of 29943 fundus photographs of 7517 patients were evaluated. In the case of a different evaluation of the degree of DR between the program and the human evaluators, the opinion of the international group of retinal specialists was decisive. Here the program attained 96.8% sensitivity (human evaluators 73.4%) and 95.6% specificity (human evaluators 98.0%) for the presence of medium and worse DR. For diabetic macular edema sensitivity was 95.3% (human evaluators 62.0%), and specificity 98.2% (hu-

man evaluators 99.2%). For severe form of DR the program attained 93.6% sensitivity (human evaluators 63.5%) and 98.2% specificity (human evaluators 99.7%).

The Inception architecture has also been used by other authors for the development of a system of artificial intelligence. Li et al. tested their program on a set of 14 520 fundus photographs and attained 92.5% sensitivity and 98.5% specificity [27]. Salhsten et al. tested their algorithm on a set of 7118 photographs and attained 89.6% sensitivity and 97.4% specificity [28].

Medios AI

Medios AI (Remidio Innovative Solutions Pvt. Ltd., Bangalore, India) is a unique system for evaluating photographs obtained by the portable ocular fundus camera Remidio Non-Mydriatic Fundus on Phone, which uses an inserted smart phone for photographing. The algorithm itself runs off-line directly on the smart phone Apple 6 (Apple Inc., Cupertino, California, USA), which is attached to the optical part of the portable fundus camera. A pilot study examined 231 patients. It was possible to perform the evaluation on 213 of these patients, and the system attained 100% sensitivity and 88.4% specificity for significant diabetic retinopathy [29].

Sosale et al. examined 900 patients [30]. The gold standard was evaluation by five retinal specialists. An evaluation was conducted of 2 non-mydriatic fundus photographs of each eye – one centred on the macula, the second on the optic nerve. In this study the system attained 83.3% sensitivity and 95.5% specificity for the presence of any kind of DR. For significant DR it attained 93.0% sensitivity and 92.5% specificity.

Another study examined 297 patients [31]. Three photographs in mydriasis were evaluated for each eye – one centred on the macula, the second from the nasal part of the retina and the third from the superotemporal region of the retina. The gold standard was evaluation by two experienced retinal specialists. The system attained 86.8% sensitivity and 95.5% specificity for the presence of any kind of DR. For significant retinopathy it attained 98.8% sensitivity and 86.7% specificity.

EyeArt

Similarly to the RetmarkerDr system, Eye Art (Eyenuk, Inc., Woodland Hills, California, USA) also was originally developed as a system for detecting signs of DR on the basis of machine learning. In this version the system demonstrated 90.0% sensitivity and 63.2% specificity in a set of 40 542 fundus photographs of 5084 patients [23]. In another study it attained 94.7% sensitivity for any kind of DR, 93.8% for significant retinopathy and 99.6% for proliferative DR. However, specificity was only 20.0% [33].

In version 2.0 the system was improved by algorithms based on deep learning and deep neural networks. The improved system was tested in a study in which it evaluated 850 908 photographs from 107 001 screening visits [34]. Of these photographs, 54 481 were obtained without mydriasis, 46 580 in mydriasis, at 649 visits the state of mydriasis was unknown. For significant DR the system attained 91.3% sensitivity and 91.1% specificity. The results were comparable for the photographs obtained without mydriasis and in mydriasis (sensitivity 89.6% and 93.0% respectively, specificity 91.7% and 94.0% respectively). Only 0.9% of visits were impossible to evaluate.

Similar results were attained by the system also in further studies. In a prospective study, Heydon et al. compared the performance of the EyeArt system with human evaluators within the framework of an English DR screening programme [35]. Photographs from 30 000 visits were evaluated within the framework of the screening programme. The system attained 95.7% sensitivity for significant DR, 100% for medium to severe non-proliferative DR and 100% for proliferative DR. The specificity of the system was 68%.

In two studies the system was tested also on photographs obtained by a portable retinal camera. Rajalakshimi et al. tested the EyeArt system on photographs obtained by the portable retinal camera Remidio Non-Mydriatic Fundus on Phone, 296 patients were evaluated. The system attained 95.8% sensitivity and 80.2% specificity for any kind of DR. For sight-threatening DR it attained 99.1% sensitivity and 80.4% specificity. Kim et al. tested the system on photographs obtained by the retinal camera RetinaScope (Retinascope, Purmerend, Netherlands). The system attained 87.0% sensitivity and 78.6% specificity. However, only 119 eyes of 69 patients were evaluated.

RetCAD

The RedCAD system (Delft Imaging Systems, 's-Hertogenbosch, Netherlands) is able to detect signs of DR and age-related macular degeneration from a photograph of the ocular fundus. The system was tested on a set of 600 photographs [36]. The performance of the system was compared with a human evaluator – optometrist, the gold standard was evaluation by four ophthalmologists. The system attained 90.1% sensitivity and 90.6% specificity for significant DR, and 91.8% and 87.5% respectively for significant age-related macular degeneration. The human evaluator attained 61.5% sensitivity and 97.8% specificity for DR, and 76.5% and 96.1% respectively for age-related macular degeneration.

Others

One of the first studies on the theme of the use of deep neural networks for detecting DR was published by Pratt et al. Their system was trained on a set of 78 000 photographs of the ocular fundus. The testing of the system then took place on 5 000 photographs, in which the system attained 30% sensitivity and 95% specificity. Gargeya and Leng used over 75 000 fundus photographs for the training of their algorithm. On two publicly available sets of photographs of the ocular fundus of patients with diabetes they attained 93.0% sensitivity and 87.0% specificity, and 90.0% and 94.0% respectively. The system was able to evaluate one photograph within 6 seconds on a regular desktop computer, and within 8 seconds on a smart phone.

Riaz et al. published the results of their algorithm, which uses deep and densely connected neural networks [37]. The system was trained on a set of 72 000 photographs and tested on two publicly available sets of photographs of the ocular fundus of patients with diabetes.

The first set contained 1747 photographs and the system attained 98.0% sensitivity and 98.0% specificity. The second tested set contained almost 18 000 fundus photographs, and the system attained 94.0% sensitivity and 97.0% specificity. This is the best result yet attained by artificial intelligence on these two sets of photographs.

DISCUSSION

The ultimate objective within the framework of diagnosis of DR is indisputably screening of the largest possible cohort of diabetics, ideally all, and the sending of those patients with DR of the determined degree to a specialised centre for stipulation of the further procedure. Upon the use of a sufficiently user-friendly retinal camera, a diabetic clinic appears to be the most suitable location of obtaining images.

In principle, the obtained images can be evaluated with the aid of two different approaches: with the aid of an ophthalmologist or trained specialist, or with the aid of specialised automated systems. Researchers from the Czech Republic are also contributing to the development of both of these directions.

A project entitled "Early Identification of Diabetic Retinopathy and Macular Edema in Patients with Type 1 or Type 2 Diabetes CZ.03.263/0.0/0.0/15_039/0008165" is being implemented by the Institute of Healthcare Information and Statistics of the Czech Republic. The ongoing pilot project of early identification of DR and macular edema is focused on a population of adult diabetics in the care of monitoring doctors, with the aim of increasing the percentage of realised ophthalmological examinations by introducing examination with the aid of a non-mydriatic camera. In contrast with the existing practice, the pilot project is introducing innovation in the performance of retinal examination of the patient directly in the clinic of the doctor responsible for monitoring of diabetics and sending of the images for central evaluation by a co-operating ophthalmologist from the Department of Ophthalmology at the Královské Vinohrady University Hospital (FNKV).

The doctor (diabetologist) obtains images of the retina (with a non-mydriatic camera) and sends the images electronically to a central pictorial documentation archive (PACS = Picture Archiving and Communication System). PACS is installed centrally on the server facilities of the Institute of Healthcare Information and Statistics (ÚZIS), and cameras from all co-operating centres are connected to the system. After saving of an image in PACS, the data is available to the ophthalmologist, who has access to the images via a web interface, thanks to which he or she can reach the examination from anywhere. The doctor performs evaluation of the image and saves the conclusion with the aid of the browser directly for the examination, where it is available to the monitoring diabetologist. The entire solution always has access rights configured so that the attending doctors can see only their own data, but at the same time the evaluating doctors can see all the data without restriction. Fig. 1.

The second alternative, namely evaluation of images with the aid of artificial intelligence, is represented in the Czech Republic by the Aireen company for example. Before describing the project, however, we shall briefly turn our attention to the term artificial intelligence.

The generally accepted definition of AI was determined by Marvin Minsky in 1967: "Artificial intelligence is the science of making machines do things that would require intelligence if done by men." However, intelligence directly imitating the human mind is not an essential condition, it may concern simpler systems. The necessary processes for the creation of such an algorithm incorporate learning the system, the application of the acquired knowledge and self-correction.

We mostly divided AI into strong and weak, or more preci-

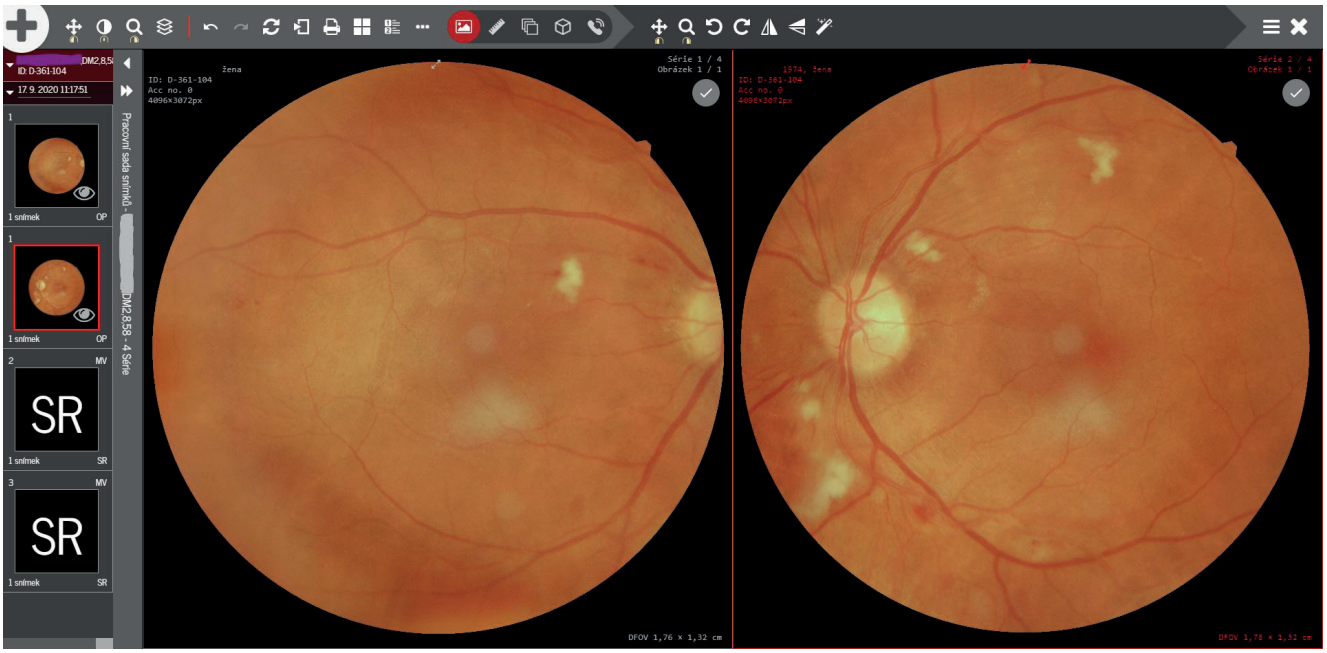

Fig. 1. Evaluation of images via web interface, source ÚZIS ČR

sely as specific and general. Specific AI mostly incorporates solutions to special tasks for which AI is trained in advance. For example, it concerns virtual assistants of the type of Siri from the Apple company or Alexa from the Amazon company. Another example was the Deep Blue system from IBM, which in 1997 defeated the chess grandmaster Garry Kasparov. These are systems which in a specific role surpass the greatest human talents, but if Deep Blue were for example to attempt to play noughts and crosses instead of chess, the machine would be useless. Such tasks are solved by systems in abstraction from connected problems, and without the interventions of their creators these systems easily fail. By contrast, general AI may be capable of acting and thinking independently, following a human model [38].

The learning process is of key importance in the creation of AI. Machine learning is a sub-discipline of AI, focusing on the creation of algorithms which, with the aid of input data, create a presupposition of the resulting situation. The situation may change over time. A classic example is advertisements on the internet. For example, if I am interested in buying a new car and I visit a number of websites on this theme, my web browser itself will offer advertisements for cars rather than handbags, until I direct my attention elsewhere. Most often, machine learning is divided into "supervised learning" and "unsupervised learning", or feedback learning. Supervised learning requires the designation of the correct result by the given expert upon entering of the input data from the beginning. In the case of unsupervised learning, the algorithms themselves seek models in the data. For example, various clusters of details, in which they reduce the dimension of data while preserving information.

Deep learning as a rule uses supervised learning – annotated images are submitted to the algorithm. Here neural networks typically appear, inspired by a simplified conception of the behaviour of neurons in the human brain. Deep learning is used for example upon image recognition or speech conversion. This type of learning requires an immense quantity of data, and in its manner is a result of the "big data" age. The disadvantage is the "black box" phenomenon, in which we are not capable of determining precisely how the system arrived at the result.

Aireen is a Czech computer program, which with the aid of machine learning enables automatic "machine" detection of specific pathological phenomena (lesions) on photographs of the retina, and thereby functions as a medical device enabling diagnosis of DR also outside an ophthalmologist's surgery. In addition to the actual diagnostic conclusion, it also provides an overview of detected pathologies, and the doctor thus has a detailed finding available, on the basis of which the conclusion was drawn. Fig. 2.

Thanks to close co-operation with experts from the ranks of Czech ophthalmologists, primarily from the Department of Ophthalmology at the Central Military Hos-

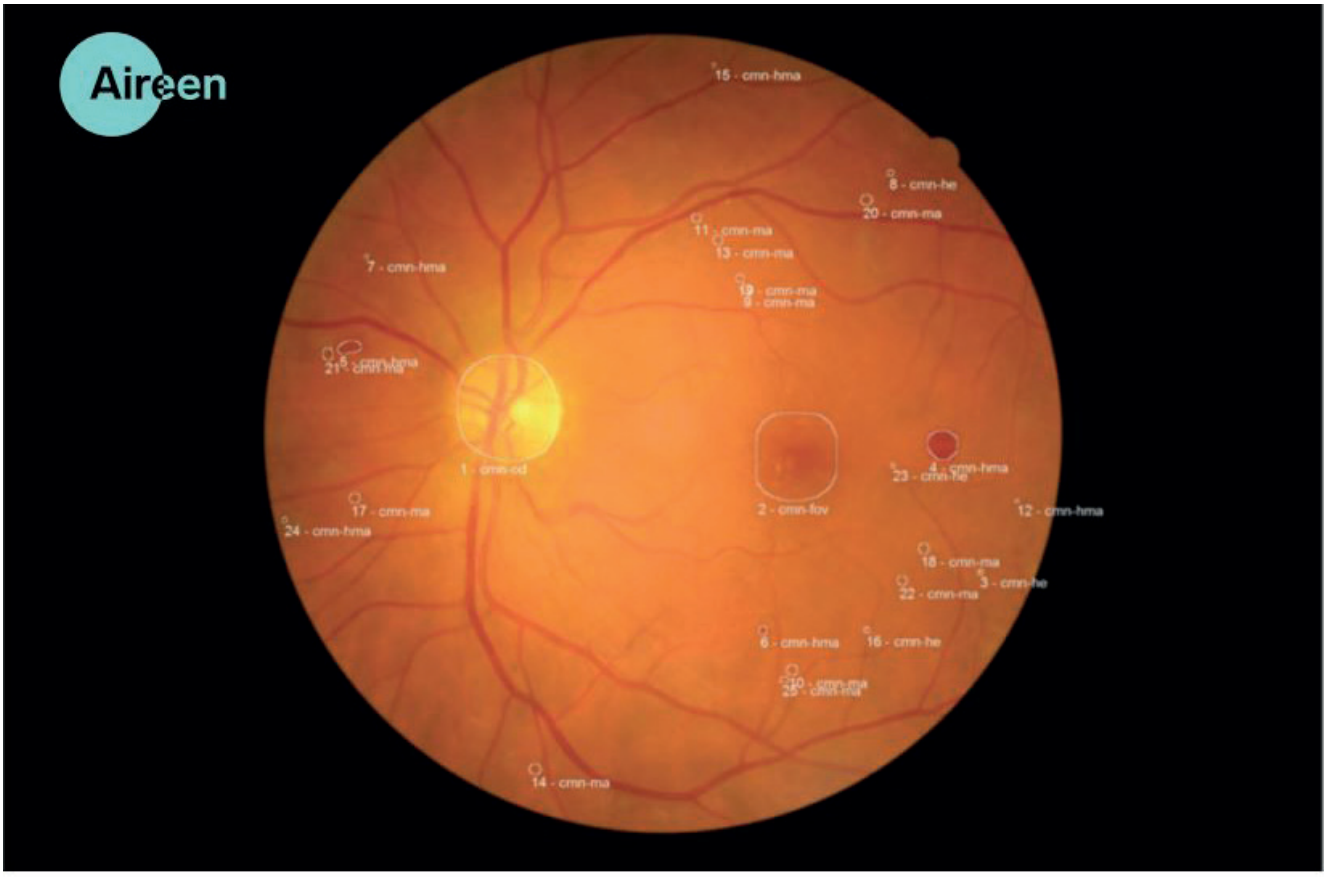

Fig. 2. Resulting evaluation with the aid of AI, source Aireen

pital (ÚVN) and other centres, Aireen has a unique model at its disposal, based on more than 12 thousand images with more than a quarter of a million designated lesions. Aireen has successfully passed a clinical trial and is now a registered medical device. The project was established two years ago, and is fully financed by Czech capital. In addition to DR, diagnosis of other conditions, including non-ophthalmological diagnoses, is in preparation.

Despite all the benefits that AI offers, it is necessary to pay attention to the problems that are linked with this technology. Today it is not computer capacity (or rather its energy demands) that is a problem, but more frequently discussed issues are of an ethical and legal nature. We are therefore moving from a theoretical and technical level to a practical level.

Michael J. Rigby reflects on the ethics of the use of AI in medicine, and notes three main challenges that need to be resolved. In part it is necessary to address the issue of balancing the risk/benefit ratio, in which the rapid introduction of AI into clinical practice may bring increased effectiveness and theoretically improve the quality of the provided care, while on the other hand issues of privacy, confidentiality (i.e. provision of data to third parties) and informed consent are not yet fully resolved. Furthermore, AI is not yet able to estimate the potential degree of co-operation that the doctor subconsciously evaluates, for example on the basis of the patient's social status, race, age, sex or overall condition, which may play a key role in deciding upon the appropriate therapy. Among other factors, these are reasons for using AI as a tool to gain a "second opinion" rather than as an element determining the ultimate diagnosis and treatment, which would entirely replace attending healthcare staff. The second theme relates to the education of doctors. There are legitimate reasons to expect that we shall be forced to shift emphasis from the knowledge--related aspect of study to extending capacities to work with AI. Already today, similar technologies are being used in the teaching of future doctors, whether they are simple instruments of the type of a simulator of indirect ophthalmoscopy, or complicated algorithms, such as a virtual patient.

The last theme is connected to the "black box" characteristic, in which we do not know how and why the algorithm arrived at the given result, which leads to problems both from the perspective of medical practice and from a legal perspective [39].

The legal status of similar algorithms differs according to individual countries, and in the Czech Republic we are waiting for the first rulings and determination of the degree of culpability in the case of failure of an AI algorithm. Nevertheless, in the majority of cases the algorithm has the status of a medical device, which can be introduced onto the European market in accordance with Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices. Before the introduction of such a medical device onto the EU market, it must successfully pass an assessment of conformity with the technical requirements stipulated in all the regulations that relate to it. The result is the issuing of a declaration of conformity and the branding of the stipulated product with the "CE" marking. A medical device introduced onto the market in accordance with this regulation must be regularly certified, registered, and the manufacturer must have a functioning system of quality control. The purpose and method of use, including the requirements for the professional qualification of the user, are defined by the manufacturer, and the product is thereby certified. The chosen manner of use has an impact on the risks that are connected with the use of the product, and which the manufacturer must safeguard against in an adequate manner. If a medical device is certified and used for the purpose stated by the manufacturer in accordance with the manner of use and in accordance with the instructions for use, the user may consider the result of the analysis of such a medical device to be reliable.

A practical multi-disciplinary problem was shown e.g. by a study of evaluating mammographic findings. In a cohort of these patients, no statistically significant difference was found in the overall evaluation of images between AI and a human evaluator, nevertheless it was demonstrated that if a human evaluated images after AI, the human had a tendency to fail in the case of images that had been evaluated by AI as falsely negative, in comparison with a situation where the human evaluated the images directly without any prior "assistance". In other words, the doctor relies excessively on the conclusion of AI and does not fulfil the role of a control mechanism if he or she has the result available in advance [40]. We have repeatedly seen a similar situation with an unclear legal framework upon testing automated vehicles, in which certain cars had difficulties with operators/drivers falling asleep at the wheel.

CONCLUSION

In the screening of diabetic retinopathy from fundus photographs, today AI surpasses its human colleagues in many aspects. The majority of systems exceed a rate of 90% specificity and sensitivity, and in the category of sight-threatening DR approach the rate of 99-100%. Nevertheless, it is necessary to be aware that for the moment AI evaluates raw data, without the connecting context, and it is clear that two identical findings in two entirely different patients may require diametrically different approaches. At the present moment, the authors of the article are looking forward to an improvement of the effectiveness of the entire system, and significant diagnostic support of the AI, which will play the role of a tool for a "second opinion" rather than the ultimate diagnostic unit, without human intervention.

LITERATURE

1. Kulkarni S, Seneviratne N, Baig MS, Khan AHA. Artificial Intelligence in Medicine: Where Are We Now? Acad Radiol. leden 2020;27(1):62–70.
2. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2. únor 2017;542(7639):115–118.
3. Witmer MT, Kiss S. Wide-field imaging of the retina. Surv Ophthalmol. duben 2013;58(2):143–154.
4. Piyasena MMPN, Gudlavalleti VSM, Gilbert C, Yip JL, Peto T, MacLeod D, et al. Development and Validation of a Diabetic Reti-

nopathy Screening Modality Using a Hand-Held Nonmydriatic Digital Retinal Camera by Physician Graders at a Tertiary-Level Medical Clinic: Protocol for a Validation Study. JMIR Res Protoc. 10. prosinec 2018;7(12):e10900.
5. Liew G, Michaelides M, Bunce C. A comparison of the causes of blindness certifications in England and Wales in working age adults (16-64 years), 1999-2000 with 2009-2010. BMJ Open. nor 2014;4(2):e004015–e004015.
6. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. The Lancet. ervenec 2010;376(9735):124–136.
7. Ústav zfravotnických informací a statistiky. ZDRAVOTNICTVÍ ČR: Stručný přehled činnosti oboru diabetologie a endokrinologie za období 2007–2017 NZIS REPORT č. K/1 (08/2018). Ústav zfravotnických informací a statistiky; 2018.
8. Scanlon PH. The English National Screening Programme for diabetic retinopathy 2003-2016. Acta Diabetol. erven 2017;54(6):515–525.
9. Nguyen HV, Tan GSW, Tapp RJ, Mital S, Ting DSW, Wong HT, et al. Cost-effectiveness of a National Telemedicine Diabetic Retinopathy Screening Program in Singapore. Ophthalmology. 1. prosinec 2016;123(12):2571–2580.
10. Tufail A, Kapetanakis VV, Salas-Vega S, Egan C, Rudisill C, Owen CG, et al. An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. Health Technol Assess Winch Engl. prosinec 2016;20(92):1–72.
11. Fleming AD, Goatman KA, Philip S, Prescott GJ, Sharp PF, Olson JA. Automated grading for diabetic retinopathy: a large-scale audit using arbitration by clinical experts. Br J Ophthalmol. 1. prosinec 2010;94(12):1606.
12. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 1. květen 2015;521(7553):436–444.
13. 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. 28. srpen 2018;1:39–39.
14. Yang WH, Zheng B, Wu MN, Zhu SJ, Fei FQ, Weng M, Zhang X, Lu PR. An Evaluation System of Fundus Photograph-Based Intelligent Diagnostic Technology for Diabetic Retinopathy and Applicability for Research. Diabetes Ther. 2019 Oct;10(5):1811-1822. doi: 10.1007/s13300-019-0652-0. Epub 2019 Jul 9. PMID: 31290125; PMCID: PMC6778552.
15 van der Heijden AA, Abramoff MD, Verbraak F, van Hecke MV, 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 (Copenh). nor 2018;96(1):63–68.
16. Verbraak FD, Abramoff MD, Bausch GCF, Klaver C, Nijpels G, Schlingemann RO, et al. Diagnostic Accuracy of a Device for the Automated Detection of Diabetic Retinopathy in a Primary Care Setting. Diabetes Care. 1. duben 2019;42(4):651.
17. Shah A, Clarida W, Amelon R, Hernaez-Ortega MC, Navea A, Morales-Olivas J, et al. Validation of Automated Screening for Referable Diabetic Retinopathy With an Autonomous Diagnostic Artificial Intelligence System in a Spanish Population. J Diabetes Sci Technol. 16. březen 2020;1932296820906212.
18. Grzybowski A, Brona P. A pilot study of autonomous artificial intelligence-based diabetic retinopathy screening in Poland. Acta Ophthalmol (Copenh) [Internet]. 3. květen 2019 [citován 6. říjen 2019];0(0). Dostupné z: https://doi.org/10.1111/aos.14132
19. Ting DSW, Cheung CY-L, Lim G, Tan GSW, Quang ND, Gan A, 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. 12. prosinec 2017;318(22):2211–2223.
20. Ribeiro ML, Nunes SG, Cunha-Vaz JG. Microaneurysm turnover at the macula predicts risk of development of clinically significant macular edema in persons with mild nonproliferative diabetic retinopathy. Diabetes Care. 2012/11/30 vyd. květen 2013;36(5):1254–1259.
21. Ramos JD, Almeida N, Oliveira CM, Neves C, Ribeiro L. 28th Meeting of the European Association for the Study of Diabetes Eye Complications Study Group (EASDec): Belfast Northern Ireland 24th – 26th May 2018. Eur J Ophthalmol. 1. květen 2018;28(1_suppl):1–38.
22. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going Deeper with Convolutions. 16. září 2014;
23. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, inception--resnet and the impact of residual connections on learning. ArXiv Prepr ArXiv160207261. 2016;
24. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 13. prosinec 2016;316(22):2402–2410.
25. Krause J, Gulshan V, Rahimy E, Karth P, Widner K, Corrado GS, et al. Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy. Ophthalmology. 1. srpen 2018;125(8):1264–1272.
26. Raumviboonsuk P, Krause J, Chotcomwongse P, Sayres R, Raman R, Widner K, et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. NPJ Digit Med. 10. duben 2019;2:25–25.
27. Li Z, Keel S, Liu C, He Y, Meng W, Scheetz J, et al. An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs. Diabetes Care. 1. prosinec 2018;41(12):2509.
28. Sahlsten J, Jaskari J, Kivinen J, Turunen L, Jaanio E, Hietala K, et al. Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading. Sci Rep. ervenec 2019;9(1):10750–10750.
29. Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S. Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening With an Offline Artificial Intelligence System on a Smartphone. JAMA Ophthalmol. 8. srpen 2019;137(10):1182–1188.
30. Sosale B, Aravind SR, Murthy H, Narayana S, Sharma U, Gowda SGV, et al. Simple, Mobile-based Artificial Intelligence Algorithm in the detection of Diabetic Retinopathy (SMART) study. BMJ Open Diabetes Res Care. leden 2020;8(1):e000892.
31. Sosale B, Sosale AR, Murthy H, Sengupta S, Naveenam M. Medios- An offline, smartphone-based artificial intelligence algorithm for the diagnosis of diabetic retinopathy. Indian J Ophthalmol. nor 2020;68(2):391–395.
32. Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala MG, Sadda S, et al. Automated Diabetic Retinopathy Screening and Monitoring Using Retinal Fundus Image Analysis. J Diabetes Sci Technol. nor 2016;10(2):254–261.
33. Tufail A, Rudisill C, Egan C, Kapetanakis VV, Salas-Vega S, Owen CG, et al. Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders. Ophthalmology. 1. březen 2017;124(3):343–351.
34. Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala MG, Sadda SR, et al. 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. 2019/08/07 vyd. listopad 2019;21(11):635–643.
35. Heydon P, Egan C, Bolter L, Chambers R, Anderson J, Aldington S, et al. Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30 000 patients. Br J Ophthalmol. erven 2020;bjophthalmol-2020-316594.
36. González-Gonzalo C, Sánchez-Gutiérrez V, Hernández-Martínez P, Contreras I, Lechanteur YT, Domanian A, et al. Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration. Acta Ophthalmol (Copenh). 2019/11/26 vyd. erven 2020;98(4):368–377.
37. Riaz H, Park J, Choi H, Kim H, Kim J. Deep and Densely Connected Networks for Classification of Diabetic Retinopathy. Diagn Basel Switz. 2. leden 2020;10(1):24.
38. Goertzel B. Artificial General Intelligence: Concept, State of the Art, and Future Prospects. J Artif Gen Intell. 13. leden 2014;0.
39. Ethical Dimensions of Using Artificial Intelligence in Health Care. AMA J Ethics. 1. únor 2019;21(2):E121–124.
40. Lehman CD, Wellman RD, Buist DSM, Kerlikowske K, Tosteson ANA, Miglioretti DL, et al. Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection. JAMA Intern Med. 1. listopad 2015;175(11):1828–1837.