Vision Research 139 (2017) 204–210

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Computational image analysis for prognosis determination in DME

Bianca S. Gerendas [a], Hrvoje Bogunovic [a], Amir Sadeghipour [a], Thomas Schlegl [a], Georg Langs [b], Sebastian M. Waldstein [a], Ursula Schmidt-Erfurth [a,*]

[a] Christian Doppler Laboratory for Ophthalmic Image Analysis, Vienna Reading Center, Department of Ophthalmology, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria
[b] Christian Doppler Laboratory for Ophthalmic Image Analysis, Computational Imaging Research Laboratory, Department of Radiology and Image-Guided Therapy, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria

ARTICLE INFO

Article history:
Received 16 January 2017
Received in revised form 10 March 2017
Accepted 14 March 2017
Available online 9 May 2017

Number of Review = 1

Keywords:
Machine learning
Large-scale data analysis
Diabetic macular edema
Random forest
Prediction
Computational image analysis

ABSTRACT

In this pilot study, we evaluated the potential of computational image analysis of optical coherence tomography (OCT) data to determine the prognosis of patients with diabetic macular edema (DME). Spectral-domain OCT scans with fully automated retinal layer segmentation and segmentation of intraretinal cystoid fluid (IRC) and subretinal fluid of 629 patients receiving anti-vascular endothelial growth factor therapy for DME in a randomized prospective clinical trial were analyzed. The results were used to define 312 potentially predictive features at three timepoints (baseline, weeks 12 and 24) for best-corrected visual acuity (BCVA) at baseline and after one year used in a random forest prediction path. Preliminarily, IRC in the outer nuclear layer in the 3-mm area around the fovea seemed to have the greatest predictive value for BCVA at baseline, and IRC and the total retinal thickness in the 3-mm area at weeks 12 and 24 for BCVA after one year. The overall model accuracy was $R^2 = 0.21/0.23$ ($p < 0.001$). The outcomes of this pilot analysis highlight the great potential of the proposed machine-learning approach for large-scale image data analysis in DME and other retinal diseases.

1. Introduction

Imaging has become a major determinant in ophthalmology since fundus photography was introduced to the community. Many multimodal imaging techniques including optical coherence tomography (OCT), fluorescein angiography, autofluorescence imaging and OCT angiography are currently available to the clinician. On the one hand, this enables ophthalmologists to gain detailed insights into human tissue on a microscopic resolution level; on the other hand, the available data cannot be manually analyzed in depth during clinical routine because the amount of data by far exceeds the capacity of healthcare providers. Furthermore, the amount of data is set to escalate as life expectancy is rising and age-related diseases are increasingly seen in ophthalmology.

Many retinal diseases compromise vision. There have been great efforts over the last decade to identify morphological biomarkers for the definition of permanent and/or reversible vision loss and for disease and/or visual acuity prediction (Garvin et al., 2008; Kapetanakis et al., 2015; Shen, Liu, & Xu, 2016; Wong & Bressler, 2016). It has been shown that intraretinal cystoid fluid (IRC) on OCT is important for prognosis of neovascular age-related macular degeneration (AMD) therapy Kapetanakis et al., 2015, a pigment epithelial detachment followed by IRC indicates a poor visual acuity prognosis for patients with AMD (Wong & Bressler, 2016) and patients with diabetic macular edema (DME) with subretinal fluid (SRF) are more likely than patients without SRF to gain vision under therapy (Shen et al., 2016).

Some of these biomarkers have been proven, others still need to be validated in larger prospective clinical trials. Most importantly, manual evaluation of these biomarkers, even in an ophthalmic reading center setting, is usually tedious and the enormous scale of imaging data provided exceeds capacities. Therefore, automatization of image data evaluation is the future in ophthalmology, as can be seen by the large number of publications in this field, for example, on the automated segmentation in OCT (Abramoff et al., 2016; Chen et al., 2012; Sophie, Lu, & Campochiaro, 2015; Yohannan et al., 2013) or automated detection of signs of diabetic retinopathy in color fundus photography (Gerendas et al., 2014; Ritter et al., 2014; Schlegl, Waldstein, Vogl, Schmidt-Erfurth, & Langs, 2015; Schmidt-Erfurth, Waldstein, Deak, Kundi, & Simader, 2015). Many of these studies used computational methods in the field of deep learning (Abramoff et al., 2016; Ritter et al., 2014;

* Corresponding author at: Department of Ophthalmology, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria.
E-mail address: ursula.schmidt-erfurth@meduniwien.ac.at (U. Schmidt-Erfurth).

http://dx.doi.org/10.1016/j.visres.2017.03.008
0042-6989/

Schlegl et al., 2015), which is described to "truly represent the brave new world in medicine" (Gerendas et al., 2014). Our study combined computational methods for the segmentation of certain features in OCT images and a random forest regression model for prediction purposes.

The aim of this pilot study was to apply a machine-learning approach to investigate the role of prognostic morphological biomarkers from imaging in a large dataset of eyes with DME. Knowledge of the role will help to find where adaptions and fine-tuning of such an existing machine-learning pipeline are needed to apply it to large-scale DME data.

2. Materials and methods

This study was a post hoc analysis of the one-year data from the Diabetic Retinopathy Clinical Research Network (DRCR.net) Protocol T study (identifier at clinicaltrials.gov: NCT01627249). All available spectral-domain (SD)-OCT scans from monthly visits were included in the analysis data set. The data set comprised 629 patients and best-corrected visual acuity (BCVA) at baseline and one year was included for preliminary correlation purposes. Imaging and functional data were prospectively collected. The main objective of the trial was to compare different intravitreal treatment options for patients with DME. The results have been published elsewhere (Schmidt-Erfurth et al., 2016). The study was conducted in compliance with the Declaration of Helsinki. Our post hoc analysis was approved by the Ethics Committee of the Medical University of Vienna. Written informed consent for inclusion into the Protocol T trial was given by each study participant.

2.1. Automated image analysis

A fully automated computational image analysis pipeline was used to process SD-OCT images. A publicly available automated segmentation algorithm based on graph theory (Iowa Reference Algorithms) was used to delineate retinal layers (Shin, Lee, Chung, & Kim, 2012). As the layer segmentation algorithm was developed for healthy retinas, only some layer interfaces were reliably segmented in DME patients. Thus, we divided the retina into three main layers representing the inner retina (IR: from inner limiting membrane to outer nuclear layer), the outer nuclear layer (ONL), and the outer retina (OR: from outer nuclear layer to retinal pigment epithelium). A voxel segmentation method based on deep learning was applied to delineate the DME-associated exudates, i.e., IRC and SRF (Abramoff et al., 2016). Examples of the segmentations of the layers obtained, if all were segmented separately, and the corresponding exudates are shown in Fig. 1. The segmentations led to 8 segmentation features (Table 1). Fig. 2 gives an example of how the thickness maps of a single case could look like if the segmentations were applied at all timepoints. Due to limitations in computational power, only baseline, week 12 and week 24 were chosen for analysis, resulting in three feature timepoints (Table 1).

Table 1
Overview of segmentation features, feature timepoints and feature regions comprising together 312 potential features for prediction.

Segmentation features	Timepoints	Regions
Total retinal thickness	Baseline	Foveal central 1 mm
Inner retina thickness	Week 12	Parafoveal 3-mm ring
Outer nuclear layer thickness	Week 24	3-mm circle
Outer retina thickness		Perifoveal 6-mm ring
Area of intraretinal cystoid fluid		6-mm circle
Volume of intraretinal cystoid fluid		Parafoveal nasal 3 mm
Area of subretinal fluid		Parafoveal superior 3 mm
Volume of subretinal fluid		Parafoveal temporal 3 mm
		Parafoveal inferior 3 mm
		Perifoveal nasal 6 mm
TOTAL: $8 \times 3 \times 13 = 312$ FEATURES		Perifoveal superior 6 mm
		Perifoveal temporal 6 mm
		Perifoveal inferior 6 mm
8	3	13

2.2. Feature extraction and predictive modeling

A summary of the retinal structure as a series of features was obtained by spatially dividing the macular retina into nine areas according to the Early Treatment Diabetic Retinopathy Study (ETDRS) grid. The mean layer thicknesses together with the area and volume of IRC and SRF were computed for each cell of the ETDRS grid. The cells included were from the foveal area (central 1 mm), 4 parafoveal (3-mm ring) as well as the 4 perifoveal (6-mm ring) areas of the nasal, temporal, superior and inferior quadrants, resulting in 13 feature regions (Fig. 3, Table 1).

Before the feature extraction, all scans of left eyes were mirrored to conform to scans of the right eye. Finally, a machine-learning approach based on the random forest regression was used to obtain a predictive model of the BCVA. The random forest was grown with 1000 trees and one third of the features were randomly sampled as candidates at each split of a tree.

2.3. Statistical analysis

The predicted values for each patient were obtained using a ten-fold cross-validation strategy. The performance of the predictive

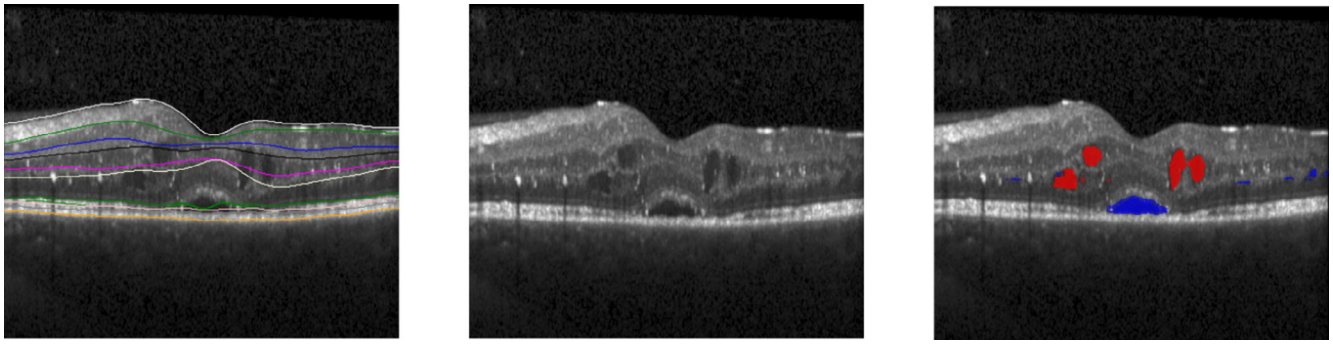

Fig. 1. Examples of the segmentations of the layers obtained (left), if all segmented were separately, and the corresponding exudates (right). Intraretinal cystoid fluid is shown in red, subretinal fluid is shown in blue. For this study, only white and green layer surfaces were used. Tissue between the top surface at the internal limiting membrane and the white surface was used as the inner retina (IR), between the white and green surface as the outer nuclear layer (ONL) and beneath the green surface down to Bruch's membrane as the outer retina (OR).

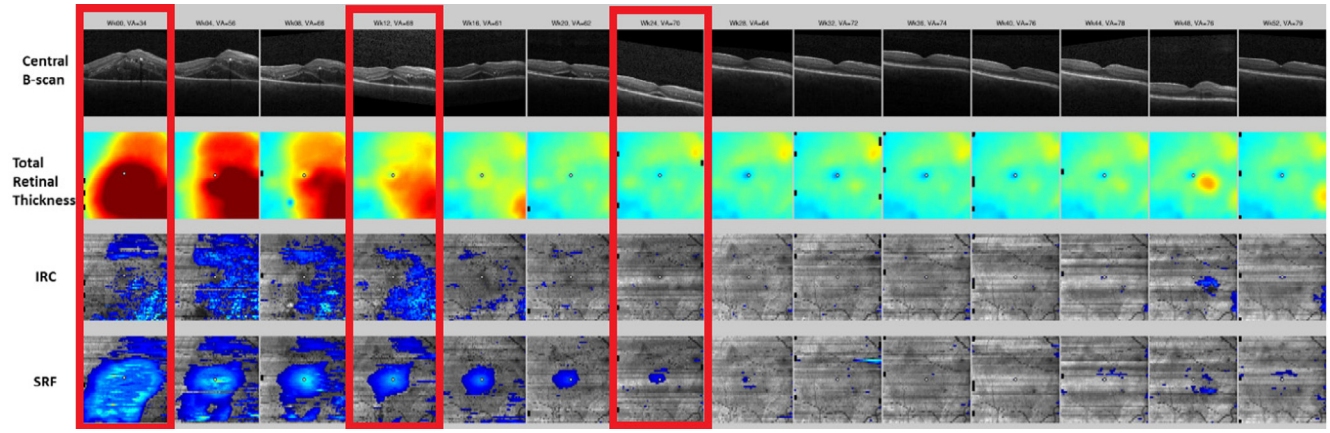

Fig. 2. Example of how the thickness maps of a single case could look like if the segmentations were applied at all timepoints. Due to limitations in computational power, only baseline, week 12 and week 24 (marked with red boxes) were chosen for analysis. From top to bottom: Central B-scan, total retinal thickness map, intraretinal cystoid fluid thickness map, subretinal fluid thickness map.

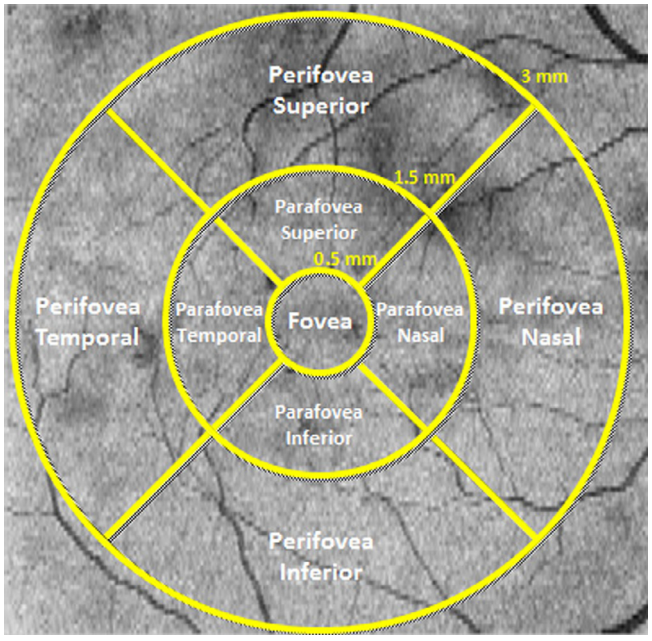

Fig. 3. Early Treatment Diabetic Retinopathy Study (ETDRS) grid. Grid areas include the foveal area (central 1 mm), 4 parafoveal as well as the 4 perifoveal areas for the nasal, temporal, superior and inferior quadrants. The 4 parafoveal areas comprise the 3-mm ring, which together with the foveal area is the 3-mm circle. The 4 perifoveal areas comprise the 6-mm ring, which together with the 3-mm circle is the 6-mm circle.

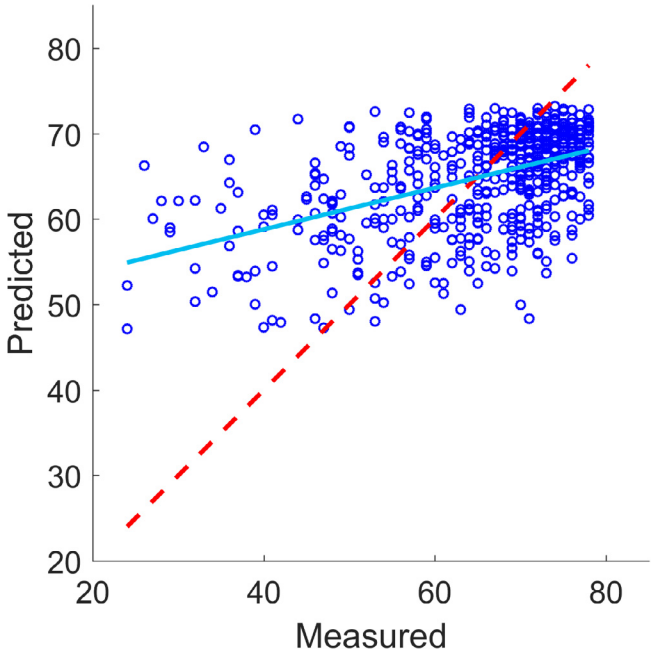

Fig. 4. Structure-function prediction using machine learning at baseline. Scatter plot of measured vs. predicted BCVA [letters]. The goodness of fit of the prediction model was $R^2 = 0.21$ ($p < 0.001$, $R = 0.46$). blue line: regression line, red line: perfect correlation regression line, where $R^2 = 1.0$.

models was evaluated by the coefficient of determination (R^2), indicating the proportion of the variance of the measured BCVA values that are explained by the predicted ones. Finally, the predictive values of the computed features were estimated by using the importance measure implemented within random forest, which relies on randomly permuting the values of a feature and measuring how such permutation affects the prediction accuracy of the model.

3. Results

3.1. OCT data and features

Of 629 patients, eight were excluded due to poor image quality (majority of B-scans missing or very low signal to noise ratio) at baseline. All remaining OCT scans (621 patients, 1863 OCT volumes) were processed and analyzed using automated segmentation. No segmentation was corrected and no wrongly segmented cases were excluded due to the exploratory character of this preliminary analysis. The number of potentially predictive features identified included in the analysis (by combining the results of different automated algorithms and applying an ETDRS grid on each scan) was 312 (104 for baseline, as only a single timepoint was used, Table 1).

3.2. Structure–function correlation at baseline

At baseline, the regression model accuracy showed $R^2 = 0.21$ ($p < 0.001$) as the level of correspondence between baseline BCVA and the proposed baseline morphologic features from OCT (Fig. 4).

Fig. 5 shows the ranking of the 20 most important features. Using our random forest model, the area of IRC in the 3-mm region around the fovea and in the 3-mm circle of the ETDRS grid seemed to have the greatest predictive value. The top 20 ranking features includes the segmentation features IRC in 5/20 (3/20 area, 2/20 vol), ONL in 5/20, total retinal thickness (TRT) in 4/20, SRF in 3/20 (2/20 vol, 1/20 area) and IR in 2/20 times. In terms of region, the center was represented in 2/20, the 3-mm ring or circle in 11/20 and the 6-mm ring or circle in 7/20 features. In more detail, the superior quadrant was represented in 3/20 cases, followed by the inferior in 2/20 and temporal/nasal each in 1/20. This finding suggests that IRC in the ONL could be important somewhere 3-mm around the fovea.

3.3. Prediction of BCVA after one year of treatment

Imaging data from baseline, week 12 and week 24 were included in the analysis to predict BCVA after one year of treatment. The regression model (Fig. 6) is able to predict the level of correspondence between BCVA after one year of treatment and the proposed morphologic features from OCT with an accuracy of $R^2 = 0.23$ ($p < 0.001$).

Fig. 7 shows the ranking of the 20 most important features with predictive value for BCVA after one year of treatment. The highest ranked feature was TRT in the central 3-mm circle around the fovea at week 24. The frequency of feature groups in the top-ranked features is shown in Table 2, demonstrating the importance of IRC and TRT in the 3-mm region after therapy was started (weeks 12 and 24 equally important).

4. Discussion

In the study presented here, an innovative machine-learning algorithm was applied to a well-defined dataset of patients with DME. Our pioneering aim was to define morphologic features on OCT at baseline and after 12 or 24 weeks of anti-vascular endothelial growth factor (VEGF) therapy that correlate well with visual acuity at baseline or after one year of therapy. The results show a fully automated ranking of the 20 most important morphologic features for the prediction of baseline BCVA and BCVA after one year. These need to be interpreted for clinical plausibility.

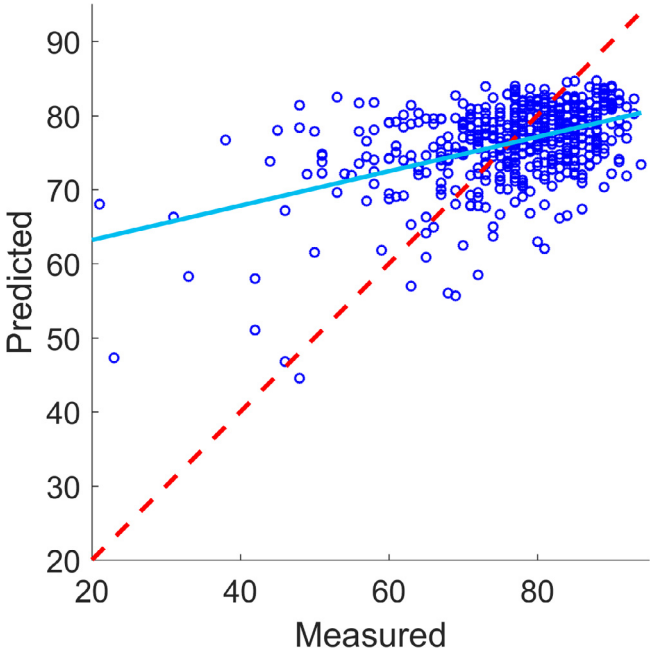

Fig. 6. Prediction of BCVA after one year of treatment using timepoints baseline, weeks 12 and 24. Scatter plot of measured vs. predicted BCVA [letters]. The goodness of fit of the prediction model was $R^2 = 0.23$ ($p < 0.001$, $R = 0.48$). blue line: regression line, red line: perfect correlation regression line, where $R^2 = 1.0$.

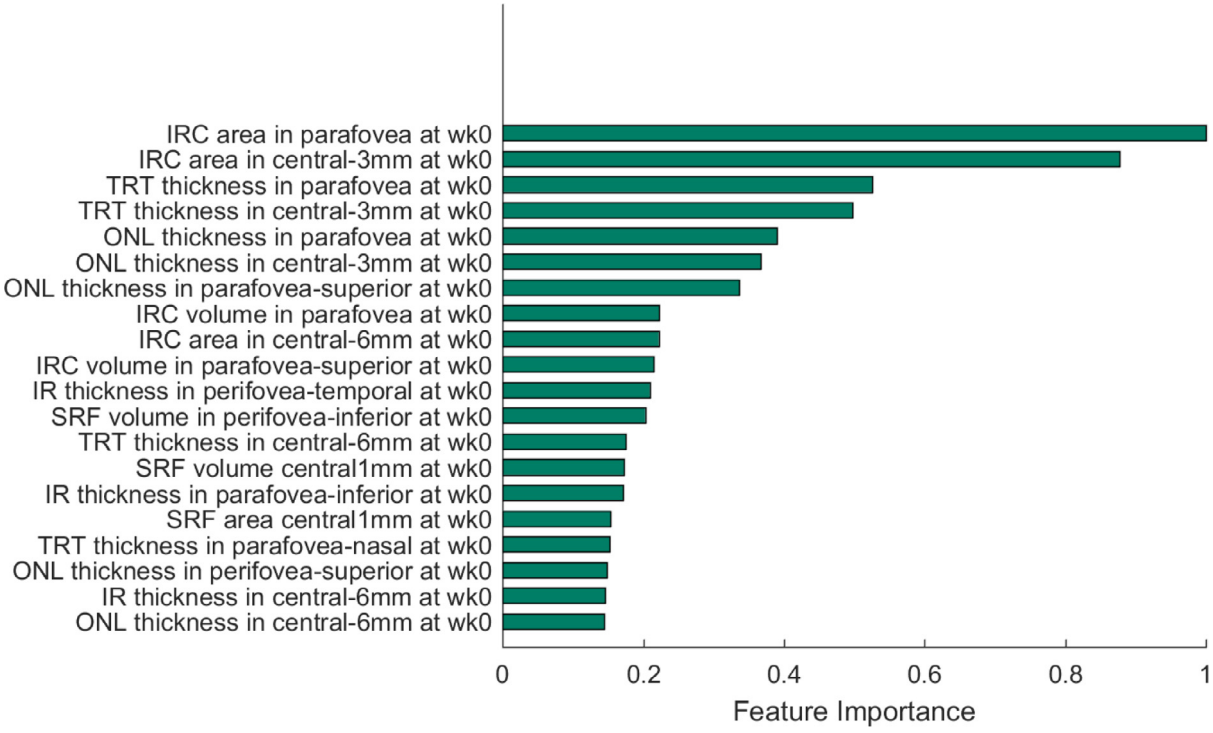

Fig. 5. Structure-function correlation using machine learning at baseline. Ranking of the top 20 features (measured normalized importance). ONL – outer nuclear layer, TRT – total retinal thickness, IR – inner retinal layers (ILM to top of ONL), SRF – subretinal fluid, IRC – intraretinal cystoid fluid.

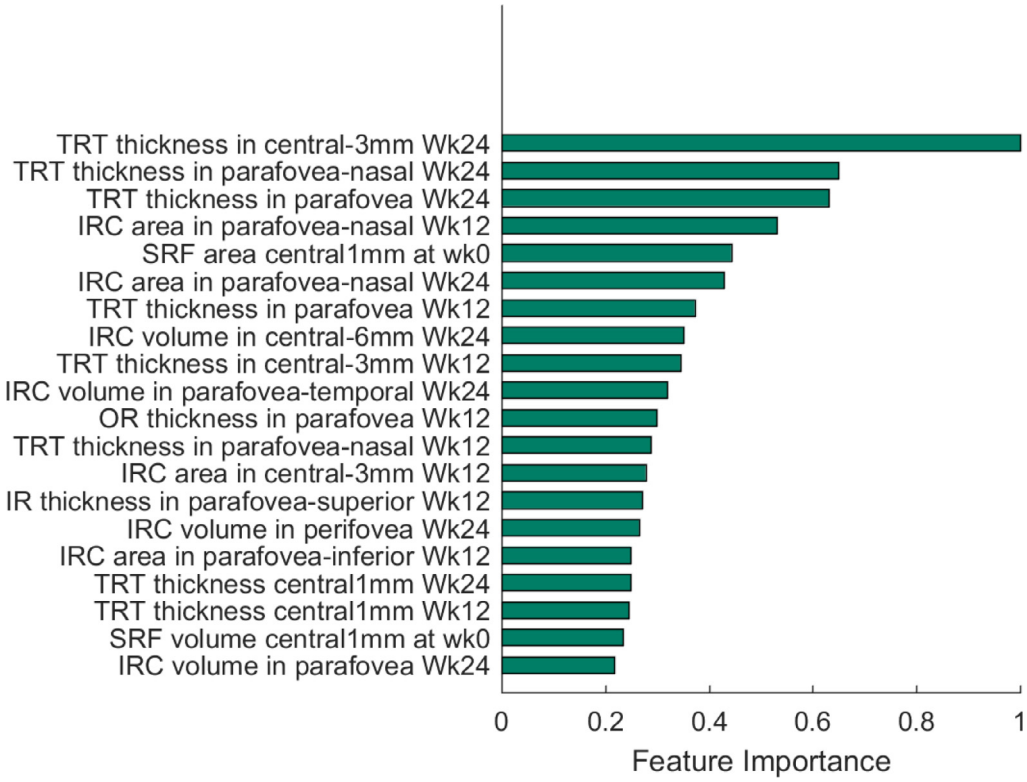

Fig. 7. Prediction of BCVA after one year of treatment using timepoints baseline, weeks 12 and 24. Ranking of the top 20 features (measured importance). TRT – total retinal thickness, ONL – outer nuclear layer, IR – inner retinal layers (ILM to top of ONL), OR – outer retinal layers (bottom of ONL to RPE), SRF – subretinal fluid, IRC – intraretinal cystoid fluid.

Table 2
Table showing top 20 ranked features according to the frequency of the appearance of single features categorized into feature, location and timepoint of importance. In brackets the frequency in 20 possible appearances is given.

Segmentation feature	Timepoint	Region
Total retinal thickness (8)	Week 12 (9)	Parafoveal 3-mm ring (11)
Intraretinal cystoid fluid (8)	Week 24 (9)	Total (4)
Area (4)	Baseline (2)	Nasal (4)
Volume (4)		Superior, temporal, inferior (each 1)
Subretinal fluid (2)		Foveal central 1 mm (4)
Area (1)		3-mm circle (3)
Volume (1)		Perifoveal 6-mm ring (1)
Outer retina thickness (1)		Total (1)
Inner retina thickness (0)		Nasal, superior, temporal, inferior (0)
Outer nuclear layer thickness (0)		6-mm circle (1)

The most important features for predicting baseline BCVA were mostly represented at the level of the ONL and the 3-mm ring and circle of the ETDRS grid. Clinically, this finding should reflect pathophysiological understanding of retinal disease. We know that the most important location for 'BCVA loss' should be in the fovea as BCVA is a functional test for the center of the retina, for a single fixation spot compared with perimetry loss where a wider field of visual perception can be tested in the periphery of the retina. Therefore, we would expect the most important features also to be located in the central 1-mm region around the fovea and not to spread to 3 mm or even be located in the 3-mm ring sparing the central 1-mm region. In addition, at first presentation patients with vision-impairing DME usually shows IRC not only in the ONL but also, in about 90% of cases, in the INL (unpublished data Prager S. et al., Medical University of Vienna).

These location variabilities may be explained by the model. The layer segmentation used was a graph-cut approach where layers are dependent on each other. As edema can destroy the layer structure with devastating consequences, the retina was divided into three main layers: The inner retina, the outer nuclear layer and the outer retina. In DME, especially at baseline, IRC spaces are often very large extending beyond anatomical layers which can make it difficult to apply a layer segmentation. Our efforts to adapt to the challenges of edematous disease by using the three main anatomic regions were probably insufficient for DME. In this condition, the border between the inner retina and outer nuclear layer is destroyed and therefore no robust segmentation can be found. If the giant IRC spaces penetrate the ONL and the INL, the anatomical borders vanish and the algorithm will usually segment around the IRC, attributing the giant IRC space to one of these layers. In our case, this was most likely in the ONL, instead of dividing the IRC and placing one half of it in the INL and the other half in the ONL. Which of these two solutions is clinically correct, remains contentious as there is literature supporting each of these solutions. Nevertheless, on a technical level, it explains the overrepresentation of the ONL and the underrepresentation of the inner retina with IRC in the INL. Clinically, it is doubtful that IRC plays a role when in the ONL but not when in the INL. But if pathophysiologically IRC first appears in the INL, it might be that the IRC in the ONL plays a larger role because it shows a more advanced disease state. In any event, most patients with vision-impairing DME already show both types of IRC at baseline.

The location to the 3-mm regions being overrepresented might be explained by the fixation alteration of patients with DME. Their fixation is usually not central and unstable when the photorecep-

tor integrity is altered, which is the case at baseline in a large number of patients with DME (Wells et al., 2015, 2016). This means that the OCT scans are not centered on the fovea and the location of the central ETDRS grid subfield can be spread to all sides because our algorithm has not centered the scans anatomically on the fovea but has set the central ETDRS grid subfield automatically to the center of the scan.

The ranking for the 20 best morphological predictors for visual acuity after one year included potential predictors from three timepoints, baseline, 12 weeks and 24 weeks, after anti-VEGF therapy induction. The explanation of the pitfalls in interpreting the baseline features easily accounts for why the later timepoints were overrepresented for prediction of BCVA after one year. The layer segmentation as well as the fluid segmentation will be more robust once the IRC spaces get smaller and are limited to a single layer or disappear. The field of the predictive features is still often 3 mm but the center is now represented more often; the center of the scan will most likely represent the fovea in more cases once the DME resolves but in many patients the timepoint might still be too early as fixation might need more time to fully recover. In some cases, fixation might never come back as the photoreceptor integrity might be permanently destroyed. Therefore, the position of the fovea still needs to be found by an algorithm or manually.

The representation of IRC at weeks 12 and 24 in this ranking is high. This could emphasize the importance of IRC behavior (response or non-response) to anti-VEGF therapy; the high representation of total retinal thickness could mirror cases where IRC has gone and the IRC area and volume is close to zero (Gulshan et al., 2016). Actually, SRF (area and volume) is the only feature represented at baseline for the prediction of BCVA after one year. This could potentially show that patients with SRF at baseline have a better potential for responding to anti-VEGF therapy and gain in vision as SRF protects the photoreceptor integrity, as has been hypothesized by our group (Gerendas et al. in submission, Philip et al. unpublished data, Medical University of Vienna) (Adhi et al., 2016; Wang et al., 2016).

Besides the above-mentioned technical improvements needed for valid clinical conclusions, our study is limited by its retrospective nature. Furthermore, baseline visual acuity, a major factor, was not included in the machine-learning algorithm for the prediction of visual acuity after one year. We know from earlier studies and also from subanalyses of this same dataset that the initial BCVA letter score of a patient is the strongest predictive factor for the final BCVA outcome; patients with good initial BCVA will have better final BCVA letter scores and patients with worse initial BCVA worse final BCVA letter scores (Philip et al., 2016; Shen et al., 2016).

In addition, other publications suggest that there are other relevant morphological features for the prediction of BCVA that we did not included in our analysis. The integrity of the photoreceptor layers, seen by a disruption of the external limiting membrane and the inner-segments/outer-segments line on OCT, seems to have a large influence (Garvin et al., 2008), which is obvious considering its direct pathophysiologic correlation to the photoreceptor function and therefore to vision. This is supported by the hypothesis, and finding, on the vision-preserving effect of SRF mentioned earlier. Although to date not part of the current algorithmic pipeline, these potential biomarkers need to be included for a clinically meaningful analysis.

Another limitation is that the patients included underwent different OCT examinations. The clinical sites in the Protocol T study used time-domain and spectral-domain OCT. Even though we only included patients who had spectral-domain OCT imaging (about 90% of the total study population), two different spectral-domain devices were used (Heidelberg Spectralis, about 60% of spectral-domain OCT examinations; Zeiss Cirrus, about 40% of spectral-domain OCT examinations). Segmentation algorithms are always stronger when used on one coherent dataset. The IRC segmentation, for example, was in the first place developed for AMD as it was trained with AMD ground-truth data. It would be needed to be trained and validated for DME on both OCT devices separately to achieve the best performance.

The final limitation is that the patients included were treated with different drugs. From the main publication of the Protocol T study (Schmidt-Erfurth et al., 2016), we know that Aflibercept leads to a better final BCVA in patients with worse baseline BCVA than bevacizumab, which is also mirrored by a better anatomical response (central retinal thickness) of the patients. This better anatomical response can be seen in both Aflibercept and Ranibizumab (Schmidt-Erfurth et al., 2016). Therefore, for clinical conclusions to be drawn our analyses should be repeated with the drug as a co-variable or with consistent patient groups only (good anatomical response, just one drug etc.).

Despite these limitations, we provide evidence that the application of our machine-learning approach has great potential for the analysis of DME datasets. Future work should aim at improving the segmentation algorithms on ground-truth data of DME patients by (1) including either manually set fovea positions or applying a validated fovea-finding algorithm, especially to the baseline scans where large edema can be seen, (2) including BCVA and photoreceptor segmentation, (3) excluding segmentation errors and (4) correcting for potential confounders such as the OCT device, baseline BCVA and drug given.

5. Conclusions

In conclusion, this study shows the immense potential of machine-learning algorithms for interpreting the big data created in modern ophthalmology. Even though earlier studies suggest that the best biomarker for the prediction of vision outcome for DME patients is baseline BCVA, machine learning has the ability to detect morphologic features and judge their predictive value better than any clinician evaluating data manually and with less bias than this conventional approach. From our preliminary analysis, we can hypothesize that IRC resolution after treatment initiation and the presence of SRF at baseline are relevant features. Nevertheless, machine learning is in its pioneering phase and procedures as well as interpretations have to be carried out with great care. If this is done properly, advanced computerized analysis of morphology will become an important guide for treatment decisions on intervals, drugs and monitoring needs. Automatically segmented features with a strong correlation to early vision, which are features beyond the current standards, will most likely be the future of morphologic OCT interpretation.

Financial disclosures

US-E: consultancy Alcon, Bayer, Boehringer Ingelheim, Novartis.
BSG: consultancy Roche.
SMW: consultancy Bayer, Novartis, research support Genentech.

Funding

This work was supported by the Christian Doppler Research Association (OPTIMA).

Data

The source of the data is the DRCR.net, but the analyses, content and conclusions presented herein are solely the responsibility of the authors and have not been reviewed or approved by DRCR.net.

References

Abramoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. C., et al. (2016). Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Investigative Ophthalmology & Visual Science, 57*(13), 5200–5206.

Adhi, M., Semy, S. K., Stein, D. W., Potter, D. M., Kuklinski, W. S., Sleeper, H. A., et al. (2016). Application of novel software algorithms to spectral-domain optical coherence tomography for automated detection of diabetic retinopathy. *Ophthalmic Surgery, Lasers and Imaging Retina, 47*(5), 410–417.

Chen, X., Niemeijer, M., Zhang, L., Lee, K., Abramoff, M. D., & Sonka, M. (2012). Three-dimensional segmentation of fluid-associated abnormalities in retinal OCT: probability constrained graph-search-graph-cut. *IEEE Transactions on Medical Imaging, 31*(8), 1521–1531.

Garvin, M. K., Abramoff, M. D., Kardon, R., Russell, S. R., Wu, X., & Sonka, M. (2008). Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search. *IEEE Transactions on Medical Imaging, 27*(10), 1495–1505.

Gerendas, B., Simader, C., Deak, G. G., Prager, S. G., Lammer, J., Waldstein, S. M., et al. (2014). Morphological parameters relevant for visual and anatomic outcomes during anti-VEGF therapy of diabetic macular edema in the RESTORE trial. *Investigative Ophthalmology & Visual Science, 55*(13), 1791.

Gerendas, B. S., Waldstein, S. M., Simader, C., Deak, G., Hajnajeeb, B., Zhang, L., et al. (2014). Three-dimensional automated choroidal volume assessment on standard spectral-domain optical coherence tomography and correlation with the level of diabetic macular edema. *American Journal of Ophthalmology, 158*(5). 1039–48 e1.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*.

Kapetanakis, V. V., Rudnicka, A. R., Liew, G., Owen, C. G., Lee, A., Louw, V., et al. (2015). A study of whether automated Diabetic Retinopathy Image Assessment could replace manual grading steps in the English National Screening Programme. *Journal of Medical Screening, 22*(3), 112–118.

Philip, A.-M., Podkowinski, D., Pablik, E., Montuoro, A., Waldstein, S. M., Gerendas, B., et al. (2016). Presence of subretinal fluid at baseline preserves from photoreceptor alterations in diabetic macular edema and cystoid macular edema due to central retinal vein occlusion. *Investigative Ophthalmology & Visual Science, 57*(12), 4164.

Ritter, M., Simader, C., Bolz, M., Deak, G. G., Mayr-Sponer, U., Sayegh, R., et al. (2014). Intraretinal cysts are the most relevant prognostic biomarker in neovascular age-related macular degeneration independent of the therapeutic strategy. *British Journal of Ophthalmology, 98*(12), 1629–1635.

Schlegl, T., Waldstein, S. M., Vogl, W. D., Schmidt-Erfurth, U., & Langs, G. (2015). Predicting semantic descriptions from medical images with convolutional neural networks. *Information Processing in Medical Imaging, 24*, 437–448.

Schmidt-Erfurth, U., Bogunovic, H., Schlegl, T., Sadeghipour, A., Waldstein, S. M., & Gerendas, B. S. (2016). Prognostic factors in the treatment of diabetic macular edema (DME) using aflibercept, ranibizumab and bevacizumab (DRCR.net protocol T). *Investigative Ophthalmology & Visual Science*. ARVO Annual Meeting 2016.

Schmidt-Erfurth, U., Waldstein, S. M., Deak, G. G., Kundi, M., & Simader, C. (2015). Pigment epithelial detachment followed by retinal cystoid degeneration leads to vision loss in treatment of neovascular age-related macular degeneration. *Ophthalmology, 122*(4), 822–832.

Shen, Y., Liu, K., & Xu, X. (2016). Correlation between visual function and photoreceptor integrity in diabetic macular edema: spectral-domain optical coherence tomography. *Current Eye Research, 41*(3), 391–399.

Shin, H. J., Lee, S. H., Chung, H., & Kim, H. C. (2012). Association between photoreceptor integrity and visual outcome in diabetic macular edema. *Graefe's Archive for Clinical and Experimental Ophthalmology, 250*(1), 61–70.

Sophie, R., Lu, N., & Campochiaro, P. A. (2015). Predictors of functional and anatomic outcomes in patients with diabetic macular edema treated with ranibizumab. *Ophthalmology*.

Wang, J., Zhang, M., Pechauer, A. D., Liu, L., Hwang, T. S., Wilson, D. J., et al. (2016). Automated volumetric segmentation of retinal fluid on optical coherence tomography. *Biomedical Optics Express, 7*(4), 1577–1589.

Wells, J. A., Glassman, A. R., Ayala, A. R., Jampol, L. M., Aiello, L. P., Antoszyk, A. N., et al. (2015). Aflibercept, bevacizumab, or ranibizumab for diabetic macular edema. *The New England Journal of Medicine, 372*(13), 1193–1203.

Wells, J. A., Glassman, A. R., Jampol, L. M., Aiello, L. P., Antoszyk, A. N., Baker, C. W., et al. (2016). Association of baseline visual acuity and retinal thickness with 1-year efficacy of aflibercept, bevacizumab, and ranibizumab for diabetic macular edema. *JAMA Ophthalmology, 134*(2), 127–134.

Wong, T. Y., & Bressler, N. M. (2016). Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA*.

Yohannan, J., Bittencourt, M., Sepah, Y. J., Hatef, E., Sophie, R., Moradi, A., et al. (2013). Association of retinal sensitivity to integrity of photoreceptor inner/outer segment junction in patients with diabetic macular edema. *Ophthalmology, 120*(6), 1254–1261.