

Impact of lens autofluorescence and opacification on retinal imaging

Leon von der Emde ,^{1,2} Geena C Rennen,¹ Marc Vaisband,^{3,4} Jan Hasenauer,^{5,6} Raffael Liegl ,¹ Monika Fleckenstein,^{1,7} Maximilian Pfau,⁸ Frank G Holz,⁹ Thomas Ach¹

To cite: von der Emde L, Rennen GC, Vaisband M, et al. Impact of lens autofluorescence and opacification on retinal imaging. *BMJ Open Ophthalmology* 2024;**9**:e001628. doi:10.1136/bmjophth-2023-001628

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjophth-2023-001628>).

Received 28 December 2023
Accepted 28 March 2024

ABSTRACT

Background Retinal imaging, including fundus autofluorescence (FAF), strongly depends on the clearness of the optical media. Lens status is crucial since the ageing lens has both light-blocking and autofluorescence (AF) properties that distort image analysis. Here, we report both lens opacification and AF metrics and the effect on automated image quality assessment.

Methods 227 subjects (range: 19–89 years old) received quantitative AF of the lens (LQAF), Scheimpflug, anterior chamber optical coherence tomography as well as blue/green FAF (BAF/GAF), and infrared (IR) imaging. LQAF values, the Pentacam Nucleus Staging score and the relative lens reflectivity were extracted to estimate lens opacification. Mean opinion scores of FAF and IR image quality were compiled by medical readers. A regression model for predicting image quality was developed using a convolutional neural network (CNN). Correlation analysis was conducted to assess the association of lens scores, with retinal image quality derived from human or CNN annotations.

Results Retinal image quality was generally high across all imaging modalities (IR (8.25 ± 1.99) > GAF > BAF (6.6 ± 3.13)). CNN image quality prediction was excellent (average mean absolute error (MAE) 0.9). Predictions were comparable to human grading. Overall, LQAF showed the highest correlation with image quality grading criteria for all imaging modalities (eg, Pearson correlation \pm CI -0.35 (-0.50 to 0.18) for BAF/LQAF). BAF image quality was most vulnerable to an increase in lenticular metrics, while IR (-0.19 (-0.38 to 0.01)) demonstrated the highest resilience.

Conclusion The use of CNN-based retinal image quality assessment achieved excellent results. The study highlights the vulnerability of BAF to lenticular remodelling. These results can aid in the development of cut-off values for clinical studies, ensuring reliable data collection for the monitoring of retinal diseases.

INTRODUCTION

Fundus autofluorescence (FAF) imaging has become an invaluable imaging technique to monitor and diagnose retinal health and disease.^{1,2} On a cellular and subcellular level, FAF mainly originates from fluorophores in the outer retina and the retinal pigment epithelium (RPE). In RPE cells, these

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The ageing lens impacts imaging through its light-blocking properties and inherent autofluorescence (AF), which can distort image analysis. There was a recognised need for more precise methods to assess the impact of lens opacification and AF on the quality of retinal images to improve diagnostics and monitoring of retinal diseases.

WHAT THIS STUDY ADDS

⇒ The study developed a convolutional neural network (CNN) model that predicts the quality of retinal images with high accuracy, comparable to human grading. It specifically highlighted that blue AF image quality is notably susceptible to changes in lens metrics, whereas infrared imaging showed the highest resilience. This study introduces quantitative measures, such as the quantitative AF of the lens and the Pentacam Nucleus Staging score, which are highly correlated with the quality of retinal imaging.
⇒ This study provides substantial insights into how lens opacification and AF metrics affect automated image quality assessment of retinal images.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The findings of this study have important implications for the future of retinal disease monitoring and research. By establishing a robust CNN-based model for retinal image quality assessment that aligns closely with human grading, this research sets the groundwork for more accurate and reliable data collection in clinical studies. The identification of specific imaging modalities that are more affected by lens changes can guide clinicians in selecting the most appropriate imaging techniques for patients with varying degrees of lens opacification. Furthermore, the development of cut-off values for lens metrics could streamline patient selection for imaging studies, ensuring that the data collected are of sufficient quality for analysis.

fluorophores are linked to intracellular granules, which show different autofluorescence (AF) properties: melanolipofuscin and lipofuscin granules with peak excitation in the short wavelength and melanin in



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Thomas Ach; Thomas.Ach@ukbonn.de

Table 1 Cohort characteristics and lens scores

Characteristic	Value
Participants, N	227
Age, years, mean, SD, (range)	60±17 (19–89)
Gender, %	
Female	46% (n=104)
Laterality, %	
Right	54% (n=123)
PNS median, (range)	3 (0–5) (au)
0, %	23%
1, %	40%
2, %	23%
3, %	7%
4, %	2%
5, %	5%
Reflectivity AC-OCT, mean, SD (range)	4.3±1.1 (2.1–6.9) (au)
LQAF mean, SD (range)	15.2±7.0 (2.6–33.5) (au)

AC-OCT, anterior chamber optical coherence tomography; LQAF, quantitative autofluorescence of the lens; PNS, Pentacam Nucleus Score.

melanosomes and melanolipofuscin granules with peak excitation in the near-infrared (IR) spectrum.^{3–5}

In clinical settings, blue (BAF) and green FAF (GAF) imaging heavily relies on the clarity of optical media, such as the cornea, the lens and the vitreous.⁶ With age, glycation products amass in the lenticular fibres, leading to increased opacity⁷ causing reduced transmission of external light to the retina. Short wavelengths are most notably affected.⁸ This phenomenon, known as Rayleigh scattering, explains that the intensity of scattered light is inversely proportional to the wavelength's fourth power (with blue having the shortest wavelength in the visual spectrum). As a consequence, many FAF studies exclude elderly patients or those with cataract.^{9–11} However, the criteria for exclusion are not defined and vary based on the investigator, making the cut-off potentially irreproducible. This poses a significant limitation, especially since many eye diseases like age-related macular degeneration predominantly affect the elderly.¹² Moreover, this practice raises concerns about patient selection in clinical research, as it systematically excludes a significant portion of the population who are most likely to be affected by these conditions.¹³

To evaluate to what extent lenticular opacification affects retinal image quality, a reliable assessment is a prerequisite. There are different approaches to quantifying lens opacity, as previously published.¹⁴ One is to determine the cataract grade clinically by slit-lamp examination using the Lens Opacities Classification System grading score.¹⁵ The cataract is classified subjectively in terms of both its severity and anatomical position, which,

however, requires clinical experience of the grader and may introduce grader bias. In contrast, there are also several lens imaging modalities that allow more objective measurements. Scheimpflug photography, in conjunction with densitometric image analysis, is able to measure the amount of light that is back-scattered from the lens.¹⁶ Further, using swept-source anterior chamber optical coherence tomography (AC-OCT) imaging, the reflectivity of the lens can be quantified.¹⁷ It is also possible to analyse the intensity of the fourth Purkinje image across different wavelengths to accurately determine lens density and spectral transmittance.¹⁸ In addition, an alternative is to use fluorophotometry deploying BAF and GAF images to measure lens transmission¹⁹ by comparing AF measures of the anterior and posterior parts of the lens. The difference in AF between both parts can be attributed to a loss of exciting and fluorescent light in the lens. Likewise, Charng and colleagues recently described a novel method to measure lens AF (LQAF) intensities by shifting the focus of the AF acquisition to the lens.²⁰ LQAF uses tools previously developed to quantify the AF of the fundus (QAF). An internal reference simultaneously captured during image acquisition enables the comparison of AF intensities across study participants and in the follow-up.

In this study, we investigated the impact of an array of lens opacification and AF measurements on qualitative and quantitative estimates of retinal image quality. Our results serve as a first step towards successful screening of patients for clinical trials where retinal image quality is pivotal. These findings underscore the critical importance of stringent criteria in study participant selection to ensure data accuracy and reliability in such trials.

METHODS

Participants

Phakic subjects were recruited at the Department of Ophthalmology, University Hospital Bonn, Germany, between January 2018 and January 2023. Inclusion criteria were age ≥18 years, no known systemic conditions or medications affecting the eye, normal retinal evaluation with no signs of retinopathy or maculopathy (as evaluated using OCT, BAF, GAF and IR) as well as willingness and ability to provide informed consent for participation in the study. Exclusion criteria for the study eye included refractive errors ≥5.00 dioptres of spherical equivalent as assessed by autorefractometry (ARK-560A; Nidek, Gamagori, Japan), a history of glaucoma or relevant anterior segment diseases with media opacities, and any history of intraocular surgery. If both eyes met the inclusion criteria, the right eye was included. Furthermore, all subjects underwent routine ophthalmological examinations, including best-corrected visual acuity, slit-lamp biomicroscopy and indirect funduscopy.

Imaging protocol

Scheimpflug images were performed using the Pentacam (Oculus, Wetzlar, Germany) with a

Table 2 Correlation of retinal image quality with lens imaging

		Focus	Illumination	Centration	Symmetry	Absence of vignetting	MOS
IR	PNS	H	-0.28*	-0.13	0.02	-0.04	0.01
	CNN	-0.40*	-0.20	-0.18	-0.18	-0.08	-0.18
AC-OCT	H	-0.16 (-0.43, 0.14)	-0.09 (-0.37, 0.20)	-0.14 (-0.42, 0.16)	-0.18 (-0.45, 0.12)	-0.09 (-0.37, 0.21)	-0.18 (-0.45, 0.12)
	CNN	-0.20 (-0.46, 0.10)	0.13 (-0.17, 0.40)	-0.00 (-0.29, 0.29)	0.04 (-0.26, 0.32)	0.01 (-0.28, 0.30)	0.01 (-0.29, 0.29)
LQAF	H	-0.26 (-0.44, -0.06)	-0.12 (-0.31, 0.08)	-0.03 (-0.23, 0.18)	-0.17 (-0.36, 0.03)	0.01 (-0.19, 0.21)	-0.19 (-0.38, 0.01)
	CNN	-0.36 (-0.52, -0.17)*	-0.21 (-0.39, -0.01)	-0.21 (-0.40, -0.01)	-0.23 (-0.42, -0.03)	-0.0 (-0.20, 0.20)	-0.28 (-0.46, -0.09)*
BAF	PNS	H	-0.41*	-0.26*	-0.11	-0.16	0.02
	CNN	-0.42*	-0.25*	-0.28*	-0.28*	-0.21*	-0.24*
AC-OCT	H	-0.31 (-0.53, -0.05)	-0.21 (-0.45, 0.05)	-0.15 (-0.40, 0.12)	-0.16 (-0.40, 0.10)	-0.17 (-0.41, 0.10)	-0.23 (-0.46, 0.04)
	CNN	-0.26 (-0.49, -0.00)	-0.05 (-0.31, 0.22)	-0.08 (-0.33, 0.19)	-0.09 (-0.34, 0.17)	0.04 (-0.23, 0.29)	-0.11 (-0.36, 0.15)
LQAF	H	-0.35 (0.50, -0.18)*	-0.19 (-0.35, -0.01)	-0.11 (-0.29, 0.07)	-0.16 (-0.33, 0.02)	-0.07 (-0.25, 0.11)	-0.20 (-0.37, -0.03)
	CNN	-0.42 (-0.55, -0.26)*	-0.24 (-0.40, -0.07)*	-0.28 (-0.44, -0.11)*	-0.25 (-0.41, -0.08)*	-0.06 (-0.24, 0.12)	-0.30 (-0.46, -0.13)*
GAF	PNS	H	-0.35*	-0.17	-0.10	-0.15	0.03
	CNN	-0.36*	-0.20	-0.15	-0.18	-0.09	-0.12
AC-OCT	H	-0.28 (-0.51, -0.02)	-0.18 (-0.43, 0.10)	-0.21 (-0.46, 0.06)	-0.15 (-0.41, 0.12)	-0.20 (-0.45, 0.07)	-0.23 (-0.47, 0.04)
	CNN	-0.17 (-0.41, 0.11)	0.08 (-0.19, 0.34)	0.08 (-0.19, 0.34)	0.15 (-0.12, 0.40)	0.13 (-0.14, 0.38)	0.04 (-0.23, 0.31)
LQAF	H	-0.31 (-0.47, -0.14)*	-0.13 (-0.30, 0.06)	-0.15 (-0.32, 0.04)	-0.15 (-0.33, 0.03)	-0.09 (-0.27, 0.09)	-0.19 (-0.36, -0.01)
	CNN	-0.32 (-0.48, -0.15)*	-0.09 (-0.27, 0.09)	-0.10 (-0.28, 0.08)	-0.05 (-0.23, 0.14)	0.10 (-0.08, 0.28)	-0.13 (-0.30, 0.06)

Pearson correlation coefficients (Spearman correlation coefficients for PNS) and 95% CIs for retinal imaging (IR, BAF and GAF) for the five grading criteria: focus, illumination, centration, symmetry and absence of vignetting, as well as the MOS with lens metrics (AC-OCT, LQAF; PNS for MOS only). Please note that for PNS, an additional Spearman correlation was applied for the five grading criteria, as it is ordinal scaled. The dashed lines separate human from CNN-based grading. Values in bold were statistically significant (p<0.05); values with p<0.01 are additionally marked with an asterisk.

AC-OCT, anterior chamber optical coherence tomography; BAF, blue autofluorescence; CNN, convolutional neural network; GAF, green autofluorescence; H, human; IR, infrared; LQAF, quantitative autofluorescence of the lens; MOS, mean opinion scores; PNS, Pentacam Nucleus Score.

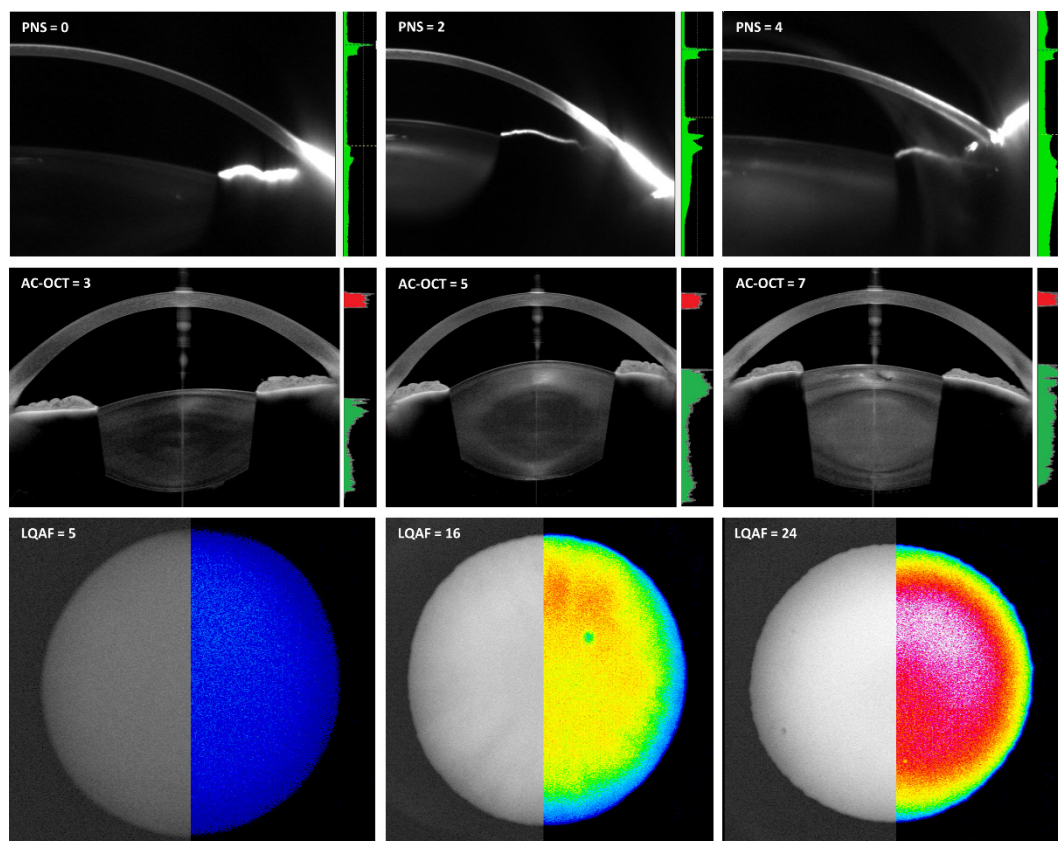


Figure 1 Lenticular imaging. Columns from left to right show three eyes with a varying degree of lenticular opacification and autofluorescence. Row one shows Scheimpflug images of three eyes with a PNS of 0, 2 and 4, respectively. On the right, a graph is shown representing the reflectivity intensity of the cornea and lens, which is provided from the manufacturer's software. Please note: the exact formula of the PNS is not published. The middle row shows AC-OCT images with averaged reflectivity values of the lens of 3, 5 and 7, respectively. The graph on the right serves to illustrate the computation of the developed AC-OCT opacification score: after normalisation, the lenticular reflectivity (in green) is compared with corneal reflectivity (in red). The lower row plots LQAF intensities, colour coded on the right. Black and blue represent low values, and red and white represent high values. PNS, Pentacam Nucleus Score; AC-OCT, anterior chamber optical coherence tomography; LQAF, quantitative autofluorescence of the lens.

standard protocol (25 images per scan, step width 10 μm).²¹ One image per subject was acquired. Following, LQAF images were performed using the Spectralis HRA (Heidelberg Engineering, Heidelberg, Germany) based on the study protocol by Charng and colleagues.²⁰ Briefly, the focus was set to +45 dioptres, and 64 images were obtained over 8 mm through the lens using the QAF mode (488 nm excitation; laser power 100%, sensitivity 67%; 30° HRA lens). Finally, subjects' lenses were imaged using a swept-source AC-OCT using default settings (Anterior Cataract App, Anterior, Heidelberg Engineering). In addition, standardised retinal imaging was performed which included combined confocal scanning laser ophthalmoscopy and spectral-domain OCT (SD-OCT) imaging (30°×25°, ART 25, 121 B-scans, Spectralis HRA-OCT 2, Heidelberg Engineering) as well as BAF (488 nm excitation, emission 500–700 nm), GAF (518 nm excitation, emission 550–700 nm) 30° FAF and IR imaging, using the same device.

Image analysis and grading

The same lenticular image analysis procedure as in our previous publication was employed. Briefly, the Pentacam Nucleus Staging (PNS) Grading score was extracted from the Scheimpflug device's software for analysis (PNS and three-dimensional cataract analysis package, Pentacam).^{22–23} PNS provides information on the mean lens density value, SD and maximum nucleus lens density and subdivides it on a scale from 0 to 5 (au) (exact formula not published by the manufacturer).

The grey values of the AC-OCT images were normalised to values between 0 and 1 using ImageJ as previously published (white=1; black=0).²⁴ In the next step, the relative reflectivity of the lens compared with the cornea was calculated. The LQAF images were imported into ImageJ as a stack of 64 bitmap images, and the LQAF was calculated according to the study protocol of Charng *et al* using the provided formula.²⁰ Briefly, the highest LQAF value from all slabs (out of 64 slabs of the z-stack) was measured in

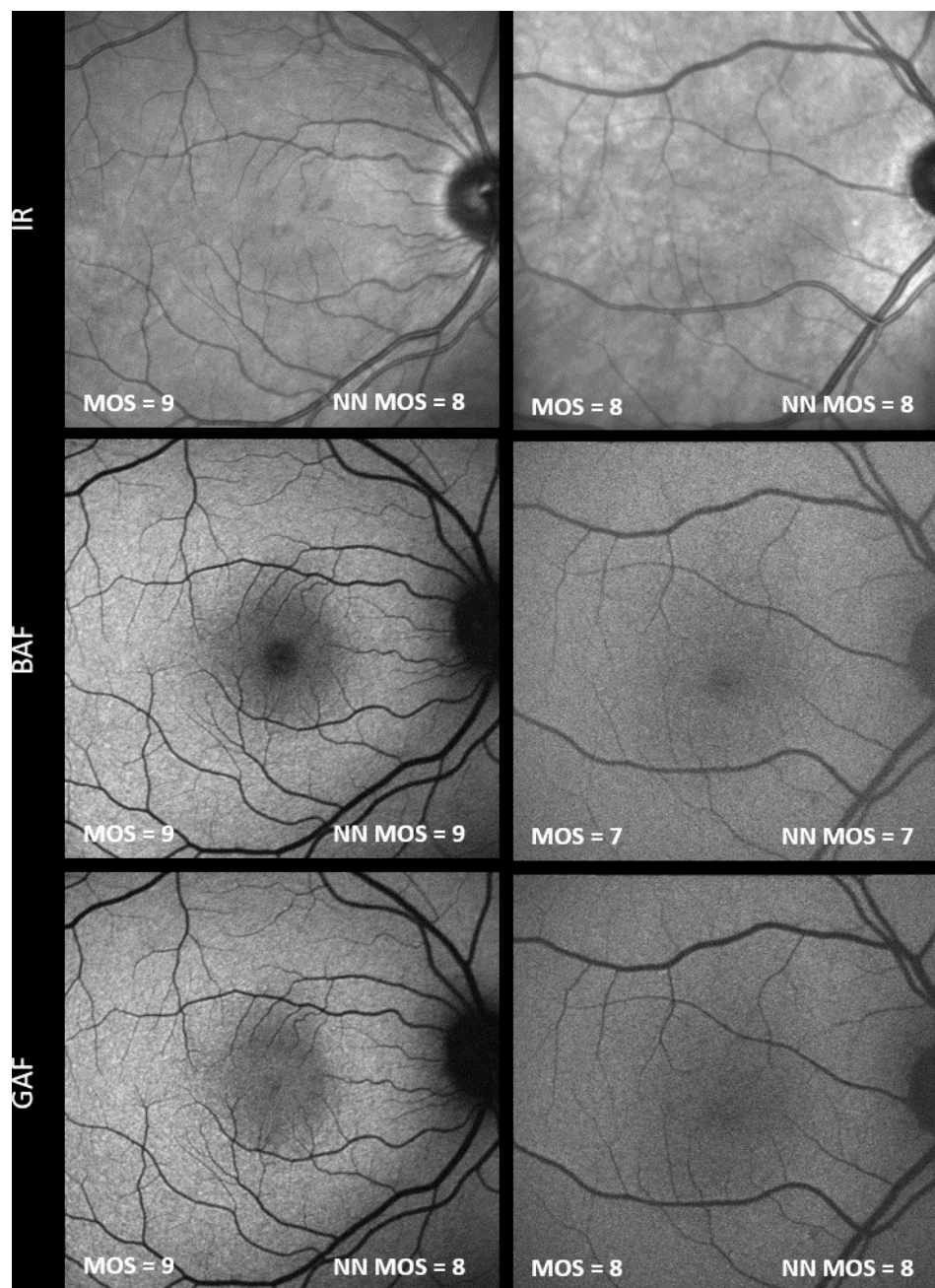


Figure 2 Multimodal retinal imaging. Retinal images with varying image quality (right column lower and left column higher). Rows one–three show IR, BAF and GAF images, respectively. In the bottom left corner of each image, the MOS (average of all five criteria and two graders) and, in the bottom right corner, CNN-based predictions are plotted. IR, infrared; BAF, blue autofluorescence; GAF, green autofluorescence; MOS, mean opinion scores; CNN, convolutional neural network.

a 60×60-pixel region in the centre of the image and divided by the AF measurement from a 200×18-pixel region of the internal reference. All retinal images (Heyex 2, Heidelberg Engineering) were exported and then subsequently graded by two readers independently (GCR and LvdE) on a scale from 1 to 10 (1 perceived as low image quality and 10 perceived as high image quality). The criteria for grading were focus (extent of small retinal vessels detectable), symmetry (upper and lower/left and right part of the image evenly captured?), illumination (image illuminated sufficiently without overexposure?), absence of

vignetting (borders of the image evenly illuminated?) and centration (is the fovea at the centre?). Following, grading was averaged to yield mean opinion scores (MOS).

Statistical analysis

For cohort characteristics, continuous variables such as age and lenticular metrics were presented using mean±SD. The distribution range for each metric was also noted. The image quality for different imaging modalities was assessed using ordinal grading. Mean values, SD and ranges were calculated for each of the grading

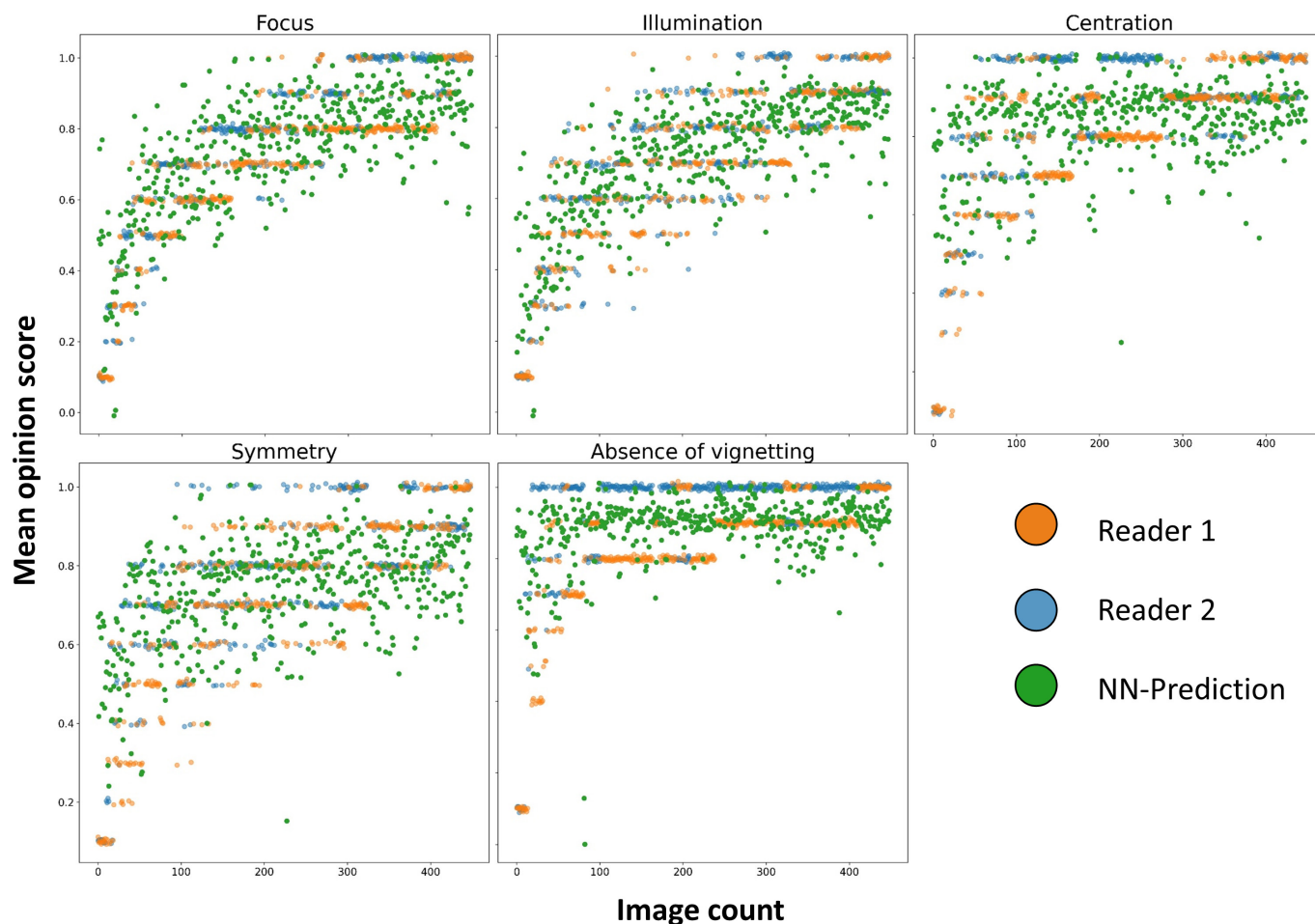


Figure 3 Image quality predictions using five different grading criteria. The y-axis shows image quality grading, and the x-axis shows the respective number of the image. Green dots represent convolutional neural network (CNN)-based grading, and blue/yellow dots represent human-based grading by the respective reader.

criteria. A convolutional neural network (CNN) architecture using three blocks of convolutions and max pooling and a fully connected regression head was implemented in the TensorFlow framework for Python (TensorFlow 2.2, Google Brain). It was trained to regress image quality with respect to the mean-squared-error (MSE) loss and evaluated using fivefold cross-validation. The details of the network architecture and training hyperparameters can be found in online supplemental figure 1. Lastly, to assess the association between retinal image quality and lens AF and opacification metrics, a correlation analysis was performed in the R programming language (R V.4.3), separately for both the human annotations and the validation results of the neural network model. Correlation values were presented with their respective CIs. In the case of non-normal distribution, no CIs were computed. The significance threshold was set at 0.05.

RESULTS

Cohort characteristics lenticular metrics

The study included 227 participants of all age groups (age range: 19–89 years; mean age \pm SD 60 \pm 17 years). Gender and eye laterality were evenly distributed (table 1). The most common PNS score was 1 (au)

with 40% followed by 0 and 2 (au) with 23% each. (table 1). The AC-OCT reflectivity was 4.3 \pm 1.1 (au) (range 2.1–6.9 (au)). The mean value for LQAF was 15.2 \pm 7.0 (au) (range 2.6–33.5 (au)) (figure 1). Our LQAF measurements were comparable with those of Charng and colleagues (see table 2C; approximately 15 (au) for age 60 years).²⁰ Similarly, our average PNS score between 1 and 2 is in line with the literature for the age group.²¹

Retinal image quality assessment

The overall image quality for all imaging modalities was high, with most metrics scoring eight and higher on average (figure 2). The mean values for all five grading criteria were highest in the IR imaging, followed by GAF and then BAF. The highest grading criteria were reached for absence of vignetting (eg, for IR images 9.4 \pm 0.8) and centration (eg, for IR images 8.2 \pm 1.4). Lower values were documented for focus with 8.4 \pm 1.4^{4–10} for IR, 7 \pm 2.3^{1–10} for BAF and 7.3 \pm 2.3^{1–10} for GAF imaging. The mean value for illumination was 8.2 \pm 1.4 (3.5–10) for IR, 7 \pm 2.3^{1–10} for BAF and 7.2 \pm 2.7 for GAF with similar image quality values for symmetry.

Automated image quality prediction

Image quality criteria predictions yielded an average MAE of 0.9 (root mean squared error [RMSE] of 1.3), with overall prediction quality being superior to the null model (MAE 1.1, RMSE 1.6). In the absence of an established null model, the (cross-validated) mean value of the target variable was used for reference. For the individual criteria, the model's performance scores were as follows (MAE/RMSE): focus (0.9, 1.2), illumination (0.8, 1.1), centration (1.0, 1.4), symmetry (1.0, 1.4) and absence of vignetting (0.7, 1.3) (figure 3). By contrast, the null model results for the individual image quality criteria were (again MAE/RMSE) as follows: focus (1.4, 1.8), illumination (1.3, 1.7), centration (0.9, 1.4), symmetry (1.1, 1.5) and absence of vignetting (0.7, 1.3). Note that the dataset was skewed towards higher image quality, especially for the criteria absence of vignetting and centration. As reflected by the near-identical performance to the null model for these criteria, the algorithm may not be as accurate for images with high centration or presence of vignetting. MOS and NN-MOS were also significantly correlated to the 'q-score' (peak signal-to-noise ratio (pSNR) provided by the manufacturer) of OCT imaging for BAF (MOS 0.395 (0.2174, 0.4867), $p<0.00001/0.1992$ (0.0464, 0.3429), $p=0.01$), GAF (0.3238 (0.1703, 0.4619), $p<0.0001/0.3067$ (0.1472, 0.4507), $p<0.001$) and IR (0.3067 (0.1472, 0.4507), $p<0.001/0.1514$ (-0.0161, 0.3107), $p<0.01$) imaging. Analysing the individual criteria, only illumination and focus demonstrated a strong correlation for all imaging modalities with the q-score.

Correlation of retinal image quality and lens metrics

Overall, LQAF showed the highest correlation with image quality grading criteria for all three imaging modalities (Pearson correlation coefficient -0.13 to -0.3, table 2), followed by PNS

and AC-OCT measurements. LQAF was particularly associated with the focus of all three imaging modalities, with highest Pearson correlation coefficient reached for BAF imaging of -0.4. The image grading criterion most correlated with lenticular metrics was focus which proved statistically significant for all retinal modalities for both LQAF and PNS (eg, -0.35 (-0.50, -0.18) for BAF/LQAF). Overall, BAF image quality showed the highest vulnerability to an increase in lenticular metrics, especially for focus, symmetry and illumination metrics. The highest resilience towards lenticular ageing proved IR imaging with only focus (eg, correlation of PNS with focus -0.28) proving statistically significant in human-based grading. CNN and human grading both yielded similar results in correlation analyses with lenticular metrics.

DISCUSSION

In this study, we examined the relationship between lenticular remodelling and retinal image quality across three retinal imaging modalities (IR, BAF and GAF). LQAF exhibited the highest correlation with image

quality grading criteria for all three imaging modalities. Notably, both neural network and human grading of image quality had comparable results in correlation analyses with lens metrics. Among the three modalities, blue AF image quality was most vulnerable to increasing lenticular opacification and AF.

We thoroughly evaluated various methods for assessing lenticular opacification and AF and its effect on retinal image quality. In our study, Charng and colleagues' measurement of LQAF was found to be both highly significant for image quality and remarkably easy to perform.²⁰ Limitations mentioned by the authors of the original study were, however, the rather small number of participants, no measurements of lenticular densitometry and no inclusion of patients with a dense cataract. With this study, we can now show that LQAF performs well in a wide age range, even in patients with dense cataracts (of note, approximately 30 patients (13 %) underwent cataract surgery within a year of study inclusion in our facility) and in patients with high lenticular opacification scores.

AC-OCT measurements of lens reflectivity were not as indicative of retinal image quality compared with other lens metrics. The swept-source OCT deployed uses a relatively long wavelength (1300nm) with a high tissue penetration.²⁵⁻²⁷ This may result in its weaker performance in assessing lenticular opacification through higher lenticular reflectivity. Nonetheless, future studies should assess different evaluation methods (eg, other normalisation techniques and varying the size of the area measurements). This holds true for the LQAF and PNS measurements as well. Mauschwitz and colleagues, when assessing the impact of lens opacity on retinal nerve fibre layer measurements, used the raw densitometry values of Scheimpflug imaging rather than the PNS.²⁸ Raw values might perform better than an ordinal-scaled score for image quality correlation as well. Additionally, the PNS is primarily a measure of nuclear cataract, and exporting raw values may have increased Scheimpflug performance on predominantly cortical cataract cases. Similarly, for LQAF measurements, we adopted the procedure by Charng and colleagues but did not assess the effect of varying the size of the area measurements of the three-dimensional lenticular AF data.²⁰ Deep learning-based methods to assess the association of lenticular ageing with retinal image quality may help to identify more suitable metrics in the future.²⁹

Our model for automated image quality assessment based on human grading showed the capabilities of machine-learning-based methods to provide accurate annotations. CNN-based image quality grading may supplement existing approaches such as material tissue contrast index, pSNR, number of details ratio and others.^{30 31} In contrast to ours, these automated assessments may not capture subtle imaging characteristics that differentiate gradable from ungradable scans.³² As our CNN assessment is trained on MOS, it may prove more sensitive in image quality variations that affect

human grading. However, for all of the above metrics to be most suitable in routine care, these metrics have to be further trained and verified on participants with retinal pathologies. An image quality assessment that would differentiate between image alterations relating to pathology (preretinal) or poor operator performance would be desirable. Here, our lens metrics may fill the void in filtering out participants with retinal image alterations due to preretinal light absorption.

Our algorithms for image quality analysis might also have a use case for tele-ophthalmic approaches. In tele-ophthalmology, image acquisition is often performed by non-physicians, which could then automatically assess image quality. In case of insufficient image quality, duplicate imaging could be performed. Including lens metrics could in the future also help identify the root cause of poor image quality (due to poor patient cooperation/ due to cataract).

Precise cut-off values for lenticular AF and opacification are critical for selecting candidate patients in clinical trials. Image quality requirements depend on the analysis performed (eg, qualitative analysis vs quantitative assessment of (quantitative) AF images).^{33–35} Our advice for AF analysis would be a MOS/CNN MOS of 7 and above (figure 2). Depending on the performed analysis, a MOS /CNN MOS of 5–6 might still be acceptable. Longitudinal studies are needed to predict the image quality decay over time for interventional studies with long study periods.

Limitations of this study include that our retinal image quality assessment did not include SD-OCT and near-IR AF imaging; examining the relationship with lens metrics could have provided additional insights.^{36 37} Given that two modalities based on AF and one modality based on reflectivity were included, it makes the comparison between modalities more difficult. Additionally, the use of both healthy and diseased retinas may have provided a more robust image quality assessment. The inclusion of objective image quality criteria such as pSNR and no reference structural similarity index may have had a higher correlation with lens metrics than the applied subjective grading. Finally, as mentioned above, alternate lens metrics and evaluation methods might have improved results.

However, the major strengths of this study are that we examined the relationship between lenticular metrics and retinal image quality across various imaging modalities in a large cohort of 227 participants across a wide age range (19–89 years). Additionally, the study incorporated both qualitative (MOS) and quantitative measures of retinal image quality, allowing for a more robust evaluation.

In conclusion, our study highlights the impact of lenticular metrics on retinal image quality and demonstrates the potential of LQAF to assess image quality before image acquisition. The strong performance of automated image quality predictions using neural networks suggests that such methods could be valuable for improving

image quality assessment in clinical practice. Especially for longitudinal studies deploying short-wavelength AF, knowledge of lenticular ageing and opacification and the effect on image quality can be valuable for patient selection. Future research should also focus on automated image quality assessment in the presence of ocular conditions.

Author affiliations

¹Department of Ophthalmology, University Hospital Bonn, Bonn, NRW, Germany

²University of Bonn, Bonn, Nordrhein-Westfalen, Germany

³Life & Medical Sciences Institute, University of Bonn, Bonn, Nordrhein-Westfalen, Germany

⁴Department of Internal Medicine III with Haematology, Laboratory for Immunological and Molecular Cancer Research, Bonn, Germany

⁵Life & Medical Sciences Institute, Bonn, Germany

⁶Helmholtz Center Munich- German Research Center for Environmental Health, Institute of Computational Biology, Bonn, Germany

⁷Department of Ophthalmology and Visual Science, University of Utah Health, Salt Lake City, Utah, USA

⁸Institute of Molecular and Clinical Ophthalmology Basel, Basel, Basel-Stadt, Switzerland

⁹Department of Ophthalmology, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany

Contributors LvdE: conceptualised the study, led the project design, supervised the entire research process, played a crucial role in the interpretation of results and critically revised the manuscript for important intellectual content. GCR: conducted the majority of the experimental work, including data collection and analysis; contributed to the writing of the manuscript, especially the methods and results section; and assisted with the design of the study's analytical framework. MV: assisted with the experimental design and was instrumental in the development and validation of the analytical methods used, played a significant role in data analysis and interpretation and contributed to manuscript revision. JH: provided expertise in statistical analysis and modelling, contributed to the design of the study's statistical approach, performed statistical analysis, assisted in interpreting the data and participated in drafting the manuscript, focusing on the results and discussion sections. RGL: involved in data collection and processing and assisted in preparing materials for the manuscript. MF: provided critical expertise in the study's subject area and contributed to the conceptualisation of the research question, played a key role in interpreting the findings and contributed to writing and revising the manuscript. MP: contributed to the study design, assisted with data collection and analysis and contributed to drafting and revising the manuscript. FGH: offered significant contributions to the conceptualisation and design of the study, provided critical feedback on the project's direction and focus, contributed to the interpretation of results and critically revised the manuscript for important intellectual content. TA: provided logistical support and facilitated access to necessary resources, assisted with data collection and management and contributed to drafting and revising the manuscript, ensuring the accuracy of data and adherence to guidelines. He also acted as a guarantor to the publication.

Funding This work was supported by the Jackstaedt Foundation to LvdE, BONFOR/SciMed Grant to TA/GR (0-137.0031), Core Grant (EY014800), and an Unrestricted Grant from Research to Prevent Blindness, New York, NY, to the Department of Ophthalmology & Visual Sciences, University of Utah to MF, and by the NIH/NEI (Grant No 1R01EY027948) to TA.

Competing interests LvdE: Heidelberg Engineering (R). GCR: BONFOR (0-137.0031, Faculty of Medicine, University of Bonn, Bonn, Germany). MV: None. JH: None. MS: Gerok Research Grant (BONFOR 0-137.0030, Faculty of Medicine, University of Bonn, Bonn, Germany), Carl Zeiss Meditec AG (F), CenterVue (F), Heidelberg Engineering (F). MP: Appellis (C), Boehringer Ingelheim (C), Daiichi Sankyo (C). MF: none. KRS: none. FGH: Acucela (C,F), Allergan (F), Apellis (C, F), Bayer (C, F), Boehringer-Ingelheim (C), Bioeq/Formycon (F,C), CenterVue (F), Ellex (F), Roche/Genentech (C,F), Geuder (C,F), Graybug (C), Gyroscope (C), Heidelberg Engineering (C,F), IvericBio (C, F), Kanghong (C,F), LinBioscience (C), NightStarX (F), Novartis (C,F), Optos (F), Oxurion (C), Pixium Vision (C,F), Oxurion (C), Stealth BioTherapeutics (C), Zeiss (F,C). TA: Apellis (C), Heidelberg Engineering (R), Roche (C), Novartis (C), Novartis (R), Bayer (C), Nidek (R).

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Ethics Committee at the University of Bonn (#305/21). All study procedures adhered to the Tenets of the Declaration of Helsinki. Before study inclusion, all participants gave written informed consent after explaining the benefits and risks of this study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data are available from the corresponding author upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Leon von der Emde <http://orcid.org/0009-0009-2630-3732>

Raffael Liegl <http://orcid.org/0000-0002-6352-6259>

REFERENCES

- von Ruckmann A, Fitzke FW, Bird AC. Distribution of fundus autofluorescence with a scanning laser Ophthalmoscope. *British Journal of Ophthalmology* 1995;79:407–12.
- Schmitz-Valckenberg S, Pfau M, Fleckenstein M, et al. Fundus autofluorescence imaging. *Prog Retin Eye Res* 2021;81:100893.
- Bermond K, von der Emde L, Tarau I-S, et al. Autofluorescent organelles within the retinal pigment epithelium in human donor eyes with and without age-related macular degeneration. *Invest Ophthalmol Vis Sci* 2022;63:23.
- Ach T, Bermond K. Autofluorescence of the human retinal pigment epithelium in normal aging and in age-related macular degeneration: Histology and clinical correlation. *Klin Monbl Augenheilkd* 2019;236:672–81.
- Bermond K, Wobbe C, Tarau I-S, et al. Autofluorescent granules of the human retinal pigment epithelium: phenotypes, intracellular distribution, and age-related topography. *Invest Ophthalmol Vis Sci* 2020;61:35.
- Reiter GS, Schwarzenbacher L, Schartmüller D, et al. Influence of lens Opacities and cataract severity on quantitative fundus autofluorescence as a secondary outcome of a randomized clinical trial. *Sci Rep* 2021;11:12685.
- Bron AJ, Vrensen GF, Koretz J, et al. The ageing lens. *Ophthalmologica* 2000;214:86–104.
- Artigas JM, Felipe A, Navea A, et al. Spectral transmission of the human crystalline lens in adult and elderly persons: color and total transmission of visible light. *Invest Ophthalmol Vis Sci* 2012;53:4076–84.
- Greenberg JP, Duncker T, Woods RL, et al. Quantitative fundus autofluorescence in healthy eyes. *Invest Ophthalmol Vis Sci* 2013;54:5684–93.
- Orellana-Rios J, Yokoyama S, Agee JM, et al. Quantitative fundus autofluorescence in non-Neovascular age-related macular degeneration. *Ophthalmic Surg Lasers Imaging Retina* 2018;49:S34–42.
- Gliem M, Müller PL, Finger RP, et al. Quantitative fundus autofluorescence in early and intermediate age-related macular degeneration. *JAMA Ophthalmol* 2016;134:817–24.
- Fleckenstein M, Keenan TD, Guymer RH, et al. Age-related macular degeneration. *Nat Rev Dis Primers* 2021;7:31.
- Chang C-Y, Yuan J, Ding S, et al. Towards fair patient-trial matching via patient-criterion level fairness constraint. *AMIA Annu Symp Proc* 2023;2023:884–93.
- von der Emde L, Rennen GC, Vaisband M, et al. Personalized lens correction improves quantitative fundus autofluorescence analysis. *Invest Ophthalmol Vis Sci* 2024;65:13.
- Gali HE, Sella R, Afshari NA. Cataract grading systems: A review of past and present. *Curr Opin Ophthalmol* 2019;30:13–8.
- Domínguez-Vicent A, Birkeldh U, Laurell C-G, et al. Objective assessment of nuclear and cortical cataracts through Scheimpflug images: agreement with the LOCS III scale. *PLoS One* 2016;11:e0149249.
- Domínguez-Vicent A, Trussardi C, Wogatai U, et al. Precision of high-resolution OCT for anterior chamber measurement: agreement with Scheimpflug imaging. *J Refract Surg* 2016;32:766–72.
- Johnson CA, Howard DL, Marshall D, et al. A noninvasive Video-based method for measuring lens transmission properties of the human eye. *Optom Vis Sci* 1993;70:944–55.
- Van Best JA. Summary of studies on the blue-green autofluorescence and light transmission of the ocular lens. *J Biomed Opt* 1996;1:243.
- Charng J, Tan R, Luu CD, et al. Imaging Lenticular autofluorescence in older subjects. *Invest Ophthalmol Vis Sci* 2017;58:4940–7.
- Pei X, Bao Y, Chen Y, et al. Correlation of lens density measured using the Pentacam Scheimpflug system with the lens Opacities classification system III grading score and visual acuity in age-related nuclear cataract. *Br J Ophthalmol* 2008;92:1471–5.
- Mayer WJ, Klaproth OK, Hengerer FH, et al. Impact of crystalline lens Opacification on effective Phacoemulsification time in Femtosecond laser-assisted cataract surgery. *Am J Ophthalmol* 2014;157:426–32.
- Nixon DR. Preoperative cataract grading by Scheimpflug imaging and effect on operative Fluidics and Phacoemulsification energy. *J Cataract Refract Surg* 2010;36:242–6.
- Ach T, Rennen G, Liegl RG, et al. Personalized lens correction for quantitative fundus autofluorescence measurements. *Invest Ophthalmol Vis Sci* 2023;64:2382.
- Potsaid B, Baumann B, Huang D, et al. Ultrahigh speed 1050nm swept source/Fourier domain OCT retinal and anterior segment imaging at 100,000 to 400,000 axial scans per second. *Opt Express* 2010;18:20029–48.
- Chan PP, Lai G, Chiu V, et al. Anterior chamber angle imaging with swept-source optical coherence tomography: comparison between CASIAII and ANTERION. *Sci Rep* 2020;10.
- Zicarelli F, Staurenghi G, Invernizzi A. Anterior segment optical coherence tomography (AS-OCT) visualization of anterior Vitritis. *Ocul Immunol Inflamm* 2023;31:1101–2.
- Mauschitz MM, Roth F, Holz FG, et al. The impact of lens Opacity on SD-OCT retinal nerve fiber layer and Bruch's membrane opening measurements using the anatomical positioning system (APS). *Invest Ophthalmol Vis Sci* 2017;58:2804–9.
- Pfau M, Walther G, von der Emde L, et al. Artificial intelligence in Ophthalmology: guidelines for physicians for the critical evaluation of studies. *Ophthalmologe* 2020;117:973–88.
- Mrak M, Grgic S, Grgic M. Picture quality measures in image compression systems. IEEE Region 8 EUROCON 2003. Computer as a Tool; 233–6. Ljubljana, Slovenia.
- Lauermann JL, Treder M, Alnawaiseh M, et al. Automated OCT angiography image quality assessment using a deep learning algorithm. *Graefes Arch Clin Exp Ophthalmol* 2019;257:1641–8.
- Wang Z, Bovik AC, Lu L. Why is image quality assessment so difficult.
- Bindewald-Wittich A, Dolar-Szczasny J, Kuenzel SH, et al. Blue-light fundus autofluorescence imaging of pigment epithelial detachments. *Eye (Lond)* 2023;37:1191–201.
- von der Emde L, Pfau M, Dysli C, et al. Artificial intelligence for morphology-based function prediction in Neovascular age-related macular degeneration. *Sci Rep* 2019;9:11132.
- Kleefeldt N, Bermond K, Tarau I-S, et al. Quantitative fundus autofluorescence: advanced analysis tools. *Transl Vis Sci Technol* 2020;9:2.
- von der Emde L, Saßmannshausen M, Morelle O, et al. Reliability of retinal layer annotation with a novel, high-resolution optical coherence tomography device: A comparative study. *Bioengineering (Basel)* 2023;10:438.
- Keilhauer CN, Delori FC. Near-infrared autofluorescence imaging of the fundus: visualization of ocular Melanin. *Invest Ophthalmol Vis Sci* 2006;47:3556–64.