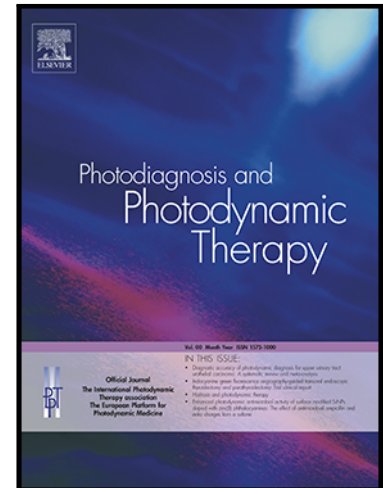


# Journal Pre-proof

Quality assessment of anterior segment OCT images: development and validation of quality criteria

Radhika Pooja Patel , Harry Petrushkin , Katie Etherton ,  
Katherine Terence , Andrew D. Dick , Jugnoo S Rahi ,  
Ameenat Lola Solebo

PII: S1572-1000(23)00613-0  
DOI: <https://doi.org/10.1016/j.pdpdt.2023.103886>  
Reference: PDPDT 103886



To appear in: *Photodiagnosis and Photodynamic Therapy*

Received date: 6 October 2023  
Revised date: 27 October 2023  
Accepted date: 3 November 2023

Please cite this article as: Radhika Pooja Patel , Harry Petrushkin , Katie Etherton , Katherine Terence , Andrew D. Dick , Jugnoo S Rahi , Ameenat Lola Solebo , Quality assessment of anterior segment OCT images: development and validation of quality criteria, *Photodiagnosis and Photodynamic Therapy* (2023), doi: <https://doi.org/10.1016/j.pdpdt.2023.103886>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Title page

**Title:** *Quality assessment of anterior segment OCT images: development and validation of quality criteria*

**Authors and affiliations:**

Radhika Pooja Patel<sup>1,2</sup>, Harry Petrushkin<sup>3,4</sup>, Katie Etherton<sup>3</sup>, Katherine Terence<sup>2</sup>, Andrew D. Dick<sup>6,7,8</sup>, Jugnoo S Rahi<sup>2,3,4,5,6,8,9,10</sup>, Ameenat Lola Solebo<sup>2,4,5,8</sup>

1. Imperial College London
2. Population, Policy and Practice Department, UCL Great Ormond Street Institute of Child Health, London, UK
3. Moorfields Eye Hospital, London, UK
4. Great Ormond Street Hospital, London, UK
5. NIHR Great Ormond Street Hospital Biomedical Research, London, UK
6. UCL Institute of Ophthalmology, London, UK
7. Bristol University, Bristol, UK
8. NIHR Moorfields Biomedical Research Centre, London, UK
9. Ulverscroft Vision Research Group, London, UK

**Corresponding author:** AL Solebo, UCL GOS Institute of Child Health, 30 Guilford Street, London WC1N 1EH; a.solebo@ucl.ac.uk; ORCID: 0000-0002-8933-5864

**Structured Abstract**

Background

The utility of medical imaging is dependent on image quality. We aimed to develop and validate quality criteria for ocular anterior segment optical coherence tomography (AS-OCT) images.

### Methods

We undertook a cross-sectional study using AS-OCT images from patients aged 6-16. A novel three-level grading system (good, limited or poor) was developed based on the presence of image artefact (categorised as lid, eyelash, cropping, glare, or movement artefact). Three independent experts graded 2825 images, with agreement assessed using confusion matrices and intraclass correlation coefficients (ICC) for each parameter.

### Results

There was very good inter-grader IQA agreement assessing image quality with ICC 0.85 (95%CI: 0.84-0.87). The most commonly occurring artefact was eyelash artefact (1008/2825 images, 36%). Graders labelled 621/2825 (22%) images as good and 384 (14%) as poor. There was complete agreement at either end of the confusion matrix with no 'good' images labelled as 'poor' by other graders, and vice versa. Similarly, there was very good agreement when assessing presence of lash (0.96,0.94-0.98), movement (0.97,0.96-0.99), glare (0.82,0.80-0.84) and cropping (0.90,0.88-0.92).

### Conclusions

The novel image quality assessment criteria (IQAC) described here have good interobserver agreement overall, and excellent agreement on the differentiation between 'good' and 'poor' quality images. The large proportion of images graded as 'limited' suggests the need for refine this classification, using the specific IQAC features, for which we also report high interobserver agreement. These findings support the future potential for wider clinical and community care implementation of AS-OCT for the diagnosis and monitoring of ocular disease.

Keywords: *optical coherence tomography; artefact; image quality*

Journal Pre-proof

## 1. Introduction

Anterior segment OCT (AS-OCT) has provided high resolution images of structure and pathology in several disease processes [1,2 ],with an emerging role in the quantification of anterior chamber inflammation, [3,4] and a possible role as a non-invasive diagnostic modality for ocular inflammation or neoplasia [5]. The clinical utility of AS-OCT, or of any medical imaging modality, is dependent on image quality. Image quality assessments interrogate and judge the properties that allow a clinician to visualise the appropriate image features necessary to come to a clinically useful conclusion [6]. These assessments may vary between clinicians [7]. Radiological studies have shown the importance of standardised, validated definitions and frameworks for image quality assessment in enabling accurate imaging-based or imaging-informed diagnosis and management [7]

Uveitis, or inflammation inside the eye, is one of the commonest reasons for attendance at eye emergency services [8] and an important cause of visual morbidity due to the structural damage caused by intraocular inflammation [9]. The disease affects all age groups, with children being at particular risk of visual loss due to asymptomatic disease [10,11]. They are also at significant risk of negative impact on quality of life and development [10–13].

Anterior uveitis (affecting the front of the eye) is graded using the presence of inflammatory cells within the normally quiet (absence of cells) anterior chamber of the eye. The traditional method of clinical grading assessment, undertaken by a specialist examining the patient using a biomicroscope (slit lamp examination, SLE) [14] is open to significant inter and intraobserver variability, and insensitive to potentially clinically impactful changes [15,16].

Anterior chamber inflammation can also be quantified through measurement of the light scattered by cells in the aqueous humour, using laser flare photometry or LFP [17]. LFP is reproducible and repeatable, but is time consuming, with poor uptake across the clinical

community [17,18]. AS-OCT is emerging as powerful tool for diagnosis and monitoring anterior uveitis, through identification and quantification of hyper-reflective particles ('cells', Fig 1) within the anterior chamber [3,15]. The absence of validated, clinically relevant image quality assessment standards for AS-OCT images has been identified by other researchers [1] and is an obstacle to the effective implementation of such imaging. We aimed to develop and validate a method to assess the quality of AS-OCT images within the context of anterior chamber inflammation quantification.

## **2. Materials and methods**

### *2.1. Subjects*

We undertook a prospective cross-sectional study anchored in a prospective longitudinal imaging biomarker validation study. Eligible participants were those aged under 18 years old under the care of a specialist uveitis service and previously diagnosed with anterior or combined anterior and intermediate uveitis. Exclusion criteria comprised the presence of corneal opacity within the central visual axis. Informed consent was obtained from all participants and guardians. Ethics Committee approval was obtained (REC reference 19/SC/0283) and the research adhered to the tenets of the Declaration of Helsinki.

### *2.2. Anterior segment imaging*

All subjects underwent swept source AS-OCT imaging using the CASIA2 (Tomey Corporation, Japan) with one of two trained specialists (KE, ALS). Acquisition protocols have been previously reported [3], but in summary involved imaging using a 64-line raster (or volume scan setting) centred at the pupil centre with settings of 12mm by 8 mm volume, 12 B-scans with 1600 A-scans per horizontal B-scan and 2 scan repeats. Images were acquired between April 2020 and December 2021.

### *3.3. Image quality grading*

Quality image criteria were defined through consensus using a three-level system. A good quality image was defined as one which allowed visualisation of the whole anterior chamber with no artefact or other image limitation. These were categorised as: (1) 'lid', the obscuration of anterior chamber visualisation due to blockage from eyelids; (2) 'lash', the presence of hyper-reflective artefactual objects in the anterior chamber due to noise from eyelash, or obscuration of anterior chamber details due to shadows from eyelashes; (3) 'glare', the presence of a hyper-reflective artefact streak within the anterior chamber due to aberration of the imaging or fixation beam; (4) 'cropping', off-centring or inadequate length of scan such that images of one or both anterior angles of the chamber was not acquired; and (5) other artefact or limiting feature (Fig 2). A poor quality image was defined (by consensus across the group) as an image artefact and / or image limitation sufficient to obscure visualisation of more than half of the central third of the anterior chamber. All images which were neither 'good' or 'poor' were defined as limited quality.

Images were anonymised, following which each B-scan underwent independent quality assessment by at least three of four graders (ALS, KE, KT, KM) who recorded the presence of each individual artefact or limiting factor, and graded the image as poor, limited and good. The graders comprised one senior ophthalmologist, one research optometrist, and two junior physicians respectively. Each of the latter graders was trained by the senior investigator (ALS) using clinical images of children known to have active disease, and graded a minimum of 100 images with an intra-grader agreement percentage of at least 98% before the study.

### 3.4. Statistical analysis

Quality assessment outcomes for each grader were analysed descriptively. Reported quality gradings and artefacts were described for each grader. In order to understand whether artefacts differed across the full volume set of images, subgroup analysis of grading outcomes across the different B scans within the volumes (divided into three groups: the upper 21, middle 21 and lower 22 B scans) was undertaken. Interobserver agreement was assessed using the Fleiss Kappa score, in order to correct for chance agreement amongst three or more graders. Very good agreement was defined as  $k > 0.8$ , good as  $k > 0.6$  to  $\leq 0.8$ , moderate as  $k > 0.4$  to  $\leq 0.6$ , and poor agreement as  $k \leq 0.4$ .<sup>19</sup> Associations between image quality and patient age were examined using univariate regression analyses with adjustment for within-child clustering. Analyses were undertaken using SPSS (18.0 software package Apache Software Foundation, Chicago, IL, USA) and Stata (Stata Statistical Software: Release 18. College Station, TX: StataCorp LLC).

## 3. Results

From 45 volume scans from 45 eyes of 37 children (23/37, 62% female, 11/37, 30% non-white ethnicity) aged 6 to 16 years, a total of 2825 B scan images were acquired for analysis. The distribution of image quality grading outcomes for these images by grader are shown in Table 1. Overall (across all graders) 612-667/2825 (22-24%) were labelled as good, and 270-388/2825 (10-14%) as poor (Table 1). Whilst using the quality grade assigned by the most senior grader as the 'gold standard', resulted in 621/2825 (22%) images being rated as good quality and 384 (14%) as poor. The most common artefact seen was the lash artefact with 958-1008/2528 images (34-37%) (Table 2). Age of participant was not associated with image quality, however, images affected by movement artefact were only acquired from patients aged under 9 years old.



### *3.1. Image quality assessment*

Assessment of the impact of horizontal scan location within the volume on scan quality showed that the lower third group had the highest proportion of images labelled as good, i.e., free of the limitations described in figure 2 (n=236- 322, 25-35%) with only 16-38 images reported as poor (1.7- 4%). The majority of the middle third group of scans were labelled limited quality and the upper third group had the highest proportion of poor images (Table 1).

Interobserver agreement of image quality assessment was very good (Fleiss kappa of 0.851,  $p < 0.001$ , CI 95% 0.835-0.867) between all three graders. There was a much higher interclass correlation coefficient in the upper third (ICC =0.823,  $p < 0.01$ , 0.794-0.852) and lower third groups (ICC =0.867,  $p < 0.01$ , 0.834-0.901) compared to the middle third group (ICC =0.798,  $p < 0.01$ , 0.766-0.830). Despite this, all three groups did have high interclass correlation coefficients consistent with strong consensus between the three graders. When assessing the correlation in a confusion matrix between graders (1 versus 2, 2 versus 3, or 3 v 1) the ICC was 1.0 with no 'good' images labelled as 'poor' by the other grader and vice versa no 'poor' images labelled as 'good'.

### *3.2. Artefact assessment*

When assessing the ability of the graders to assess the presence of artefact in the form of lashes the interclass correlation coefficient was 0.959 ( $p < 0.001$ , CI 95% 0.938-0.980) between all three graders (Table 2). When assessing the presence of lashes depending on location based on the three groups (upper, middle and lower third), there was a similar correlation in all three groups with the highest correlation in the middle third group

(ICC=0.978,  $p<0.01$  0.942-1.013). The upper third group contained the highest number of images with the presence of lash artefact ( $n=555-569$ ) compared to the middle third ( $n=168-170$ ) and the lower third group ( $n=233-271$ ). A higher number of images of left eyes were cropped on the left side (103-156 cropped right versus 271-334 cropped left) and in images of the left eye, the majority of images were cropped on the right hand side (787-887 cropped right compared to 22-30 cropped left) (Table 3).

There was similar very good agreement on assessment of glare with interclass correlation coefficient of 0.816 ( $p=0.011$ , CI 95% 0.795-0.838). The highest amount of glare was seen in the middle group ( $n=204-329$  images) compared to the upper third group where only 16-23 images were labelled as having glare and the lower third group with even fewer images ( $n=8-10$ ). Movement as an artefact had the highest level of agreement with interclass correlation coefficient of 0.972 ( $p<0.001$ , 95% CI 0.955-0.989). The most amount of movement was seen in the upper third group with between 63-67 images being reported as having movement as opposed to 15 in the middle third and between 17-20 in the lower third group.

There was very good agreement when assessing the cropping in the images both on the left (ICC=0.891,  $p<0.01$ , 0.870-0.913) and the right ICC=0.899,  $p<0.01$ , 0.878-0.920).

#### **4. Discussion**

From this prospective cross-sectional study, we report the development and use of a novel AS-OCT image quality assessment scheme and report good interobserver agreement when using it. We show that certain limiting features are clustered at different locations within volume scan sets and report reasonable agreement with regards to artefact or quality limiting

feature, with high levels of agreement on lash and movement artefact, and moderate agreement on glare artefact.

Our study findings are strengthened by a large sample size, multiple scans and artefacts and the presence of independent graders. We had multiple observers with standardised definitions of quality categories that ensured uniformity and clarity for graders. Study limitations included use only of data from children: it may be that scans acquired from adults would have a lower prevalence of artefact, or artefact of a different nature due for example to less abundant eyelashes [20]. It is likely that image quality issues are more prevalent amongst images acquired from young children, when compared to those acquired from older individuals more able to comply with direction from imaging technicians or clinicians. However, as there is no reason to presume that the artefacts and image quality factors and characteristics found in images of scans from children would not be present in images from adult, this limitation is unlikely to impact on generalisability across different age groups.

Our grading scheme is anchored in the clinical value of using the images to detect or quantify uveitis, which may make this IQA approach inappropriate for different clinical contexts without the addition of more quality characteristics. Uveitis is an important clinical entity with an active clinical research agenda, with multiple investigators currently developing pathways and tools for the automated analysis of images of anterior chamber inflammatory change [5,21–23]. However, use of these images to assess anterior chamber status necessitates unobstructed views of the structures which ‘frame’ the chamber. Consequently, our grading scheme does include quality limiting features important in different disease areas: image cropping limits utility for glaucoma, where visibility of

anterior chamber (AC) angle structures and quantification of AC volume are needed, and lash motion and cropping limit use for detection or monitoring of corneal defects. A possible critique of our grading schema is that it generated a heterogeneous category of 'limited' scans, which may benefit from further division using the degree of limitation, although our labelling of the presence of the individual quality limiting criteria does add greater granularity to the grading.

While there have not been any similar studies assessing the quality of OCT scans or investigating their interobserver agreement in ophthalmology, other arms of imaging and radiology have investigated extensively interobserver agreement and the need for quality assurance alongside understanding of the factors affecting quality. [6,7]. In order to develop a surveillance service for children at risk of uveitis, there is a need for clearly defined IQA such as exists for the national diabetic eye screening programme [24]. The disease burden of diabetes is significant and 'manual' screening by detailed clinical examination for diabetic retinopathy is impractical, therefore clear IQA for retinal images have been developed in order to create effective programmes that are accepted internationally. Our study is one of the first steps in developing a similar protocol for the use of AS-OCT in the community.

Our findings of a much higher interclass correlation coefficient in the upper third and lower third groups compared to the middle suggest it was easier to assess the quality of images at the extremes, with the highest agreement in the lower third group. The lower third group also had the highest number of images rated as 'good' by all three graders than any other group. Therefore, the group with the highest quality of images marked as 'good' also had the best correlation between the graders. Similar findings of a greater consensus on what comprises a 'good' image in validated grading schemes have been reported by investigators

working IQA for non-ocular images, for example paediatric cerebral CT images [25] and prostate MRIs [26] where images with higher quality scores had higher inter-reader agreement, and provided greater facilitation of clinical decision-making or diagnosis. The discrepancy in image quality between the three positional groups in our study may be explained by the increased presence of upper lid lash artefact that we would expect to be present more in this location.

The high interclass correlation scores of each artefact showed that these were easily identified and agreed upon by our graders and therefore clearly defined. Movement artefact was most easily identified, followed by the presence of lashes then cropping and lastly glare. When assessing cropping as expected on images of the right eye, there was a much higher number of images that were cropped on the left side and when investigating images of the left eye, majority of images were cropped on the right hand side which is likely due to position of the nose which should be taken into account when acquiring images.

Our study also highlighted the particular need for quality assessment when a raster or volume of cross sectional images is used to build a 3D representation, rather than relying solely on single AS-OCT slices in order to assess the anterior chamber of the eye. Images taken at different levels have shown varying degrees of quality and artefact can all significantly affect the image analysis when assessing cells in the anterior chamber. In order to support utility for specific clinical indications, future investigators may need to develop and refine other image quality criteria. Whilst the criteria described here are 'disease agnostic', for clinical utility for uveitis diagnosis and monitoring, there may be benefit in assessing other criteria. These include patterns in image speckling (the background image noise in anterior chamber images, potentially caused by scatter from aqueous humour

proteins or bodies). Such quality criteria may further support implementation of OCT as a widespread screening and monitoring tool for the population at risk. Other future work includes the automation of quality assessment for these and similar images, to support wider clinical and community health adoption. Such work would require large, labelled image datasets, additional to the work dataset reported here.

## **5. Conclusion**

In conclusion, this work shows that it is possible to create a valid and reproducible clinically anchored method to standardise the quality of AS-OCT images in children with uveitis. This provides the first step in developing criteria for clear IQA that can be used for future work by external groups assessing the use of AS-OCT scans in the context of anterior segment inflammation, or in other disorders involving the anterior ocular structures. Further work on refining the IQA criteria will support future implementation and support the clinical utility of AS-OCT for disease diagnosis and monitoring.

## **Declaration of Conflict of Interest**

No conflicting relationship exists for any author.

## **Acknowledgement**

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Financial disclosure(s): None of the authors have proprietary or commercial interests in any

of the materials discussed in this article.

### Figure legends

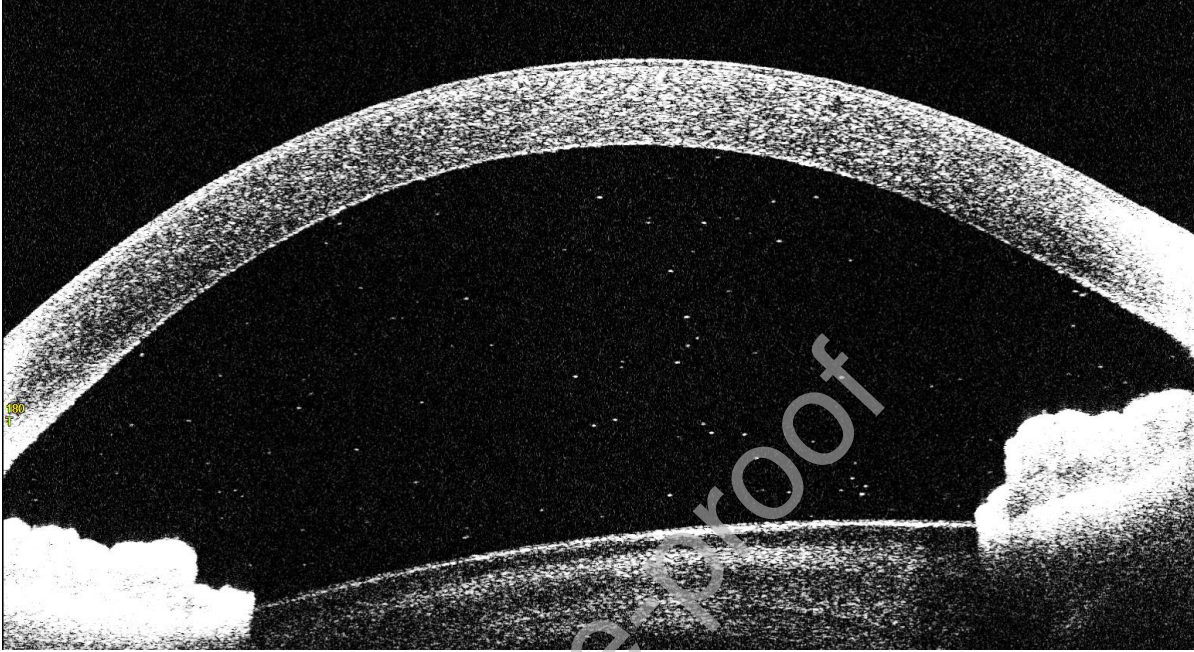


Figure 1 – Anterior Segment OCT CASIA2 (Tomey Corporation, Japan) showing presence of cells in anterior chamber (visualized as hyper-reflective specks in anterior chamber). The image has been cropped to better allow visualisation of cells

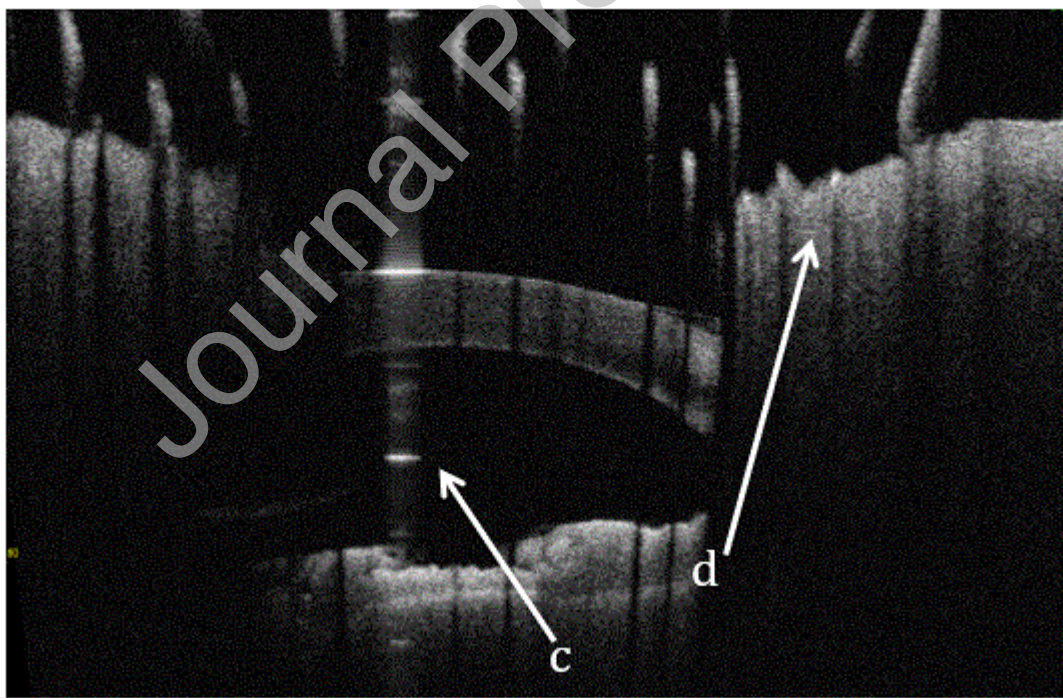
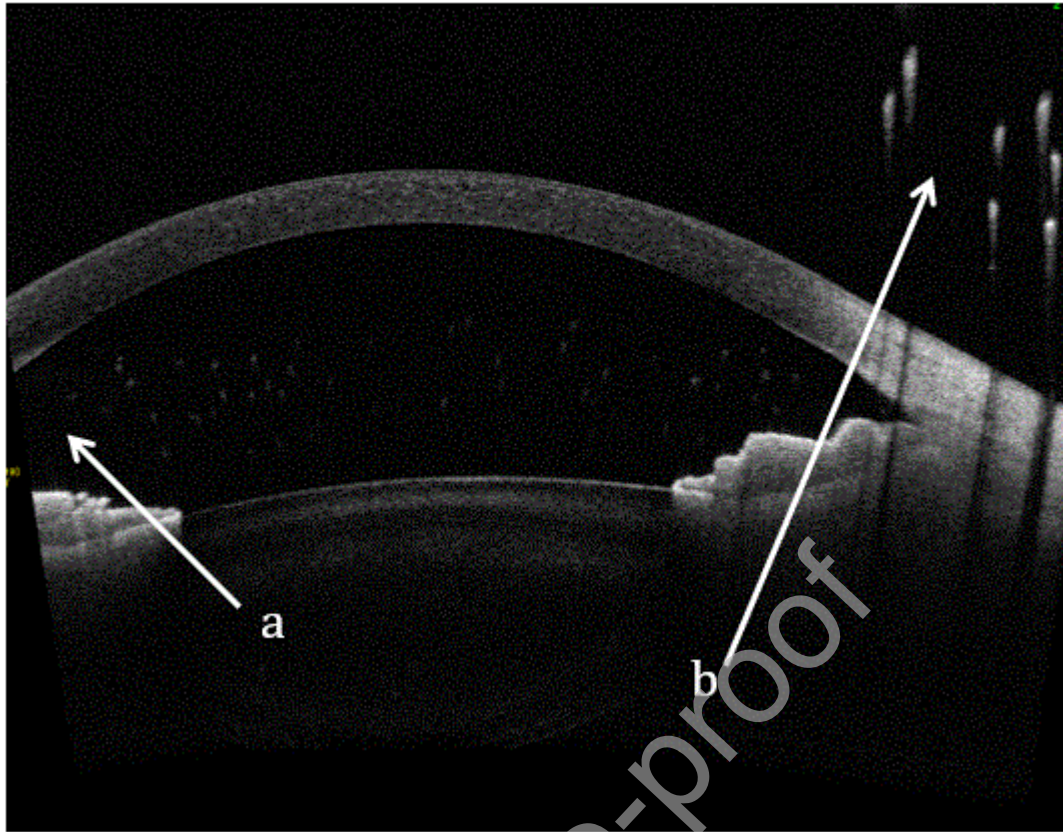


Figure 2: Showing types of artefact seen on AS-OCT a) cropping b) eyelash c) glare d) lid



## References

1. Mirzayev I, Gündüz AK, Aydın Ellialtıođlu P, Gündüz ÖÖ. Clinical applications of anterior segment swept-source optical coherence tomography: A systematic review. *Photodiagnosis and photodynamic therapy*. 2023;42:103334. doi:10.1016/j.pdpdt.2023.103334
2. Shan J, DeBoer C, Xu BY. Anterior segment optical coherence tomography: Applications for clinical care and scientific research. *Asia-Pacific Journal of Ophthalmology*. Published online 2019. doi:10.22608/APO.201910
3. Etherton K, Rahi JS, Petrushkin H, et al. Quantitative and qualitative assessment of anterior segment optical coherence tomography capture of disease state in childhood anterior uveitis. *British Journal of Ophthalmology*. Published online 2022. doi:10.1136/bjophthalmol-2021-320448
4. Akbarali S, Rahi JS, Dick AD, et al. Imaging-Based Uveitis Surveillance in Juvenile Idiopathic Arthritis: Feasibility, Acceptability, and Diagnostic Performance. *Arthritis and Rheumatology*. Published online 2021. doi:10.1002/art.41530
5. Gozzi F, Bertolini M, Gentile P, et al. Artificial Intelligence-Assisted Processing of Anterior Segment OCT Images in the Diagnosis of Vitreoretinal Lymphoma. *Diagnostics (Basel, Switzerland)*. 2023;13(14). doi:10.3390/diagnostics13142451
6. Tompe A, Sargar K. *X-Ray Image Quality Assurance*.; 2021.
7. Kjelle E, Chilanga C. The assessment of image quality and diagnostic value in X-ray images: a survey on radiographers' reasons for rejecting images. *Insights Imaging*. Published online 2022. doi:10.1186/s13244-022-01169-9
8. MacIsaac JC, Naroo SA, Rumney NJ. Analysis of UK eye casualty presentations. *Clinical & experimental optometry*. 2022;105(4):428-434. doi:10.1080/08164622.2021.1937949

9. Jones NP, Pradeep A, John B. Certification of Vision Impairment in Patients with Uveitis Attending a Specialist Clinic. *Ocular immunology and inflammation*. 2019;27(4):540-544. doi:10.1080/09273948.2018.1441874
10. Siiskonen M, Hirn I, Pesälä R, Hautala T, Ohtonen P, Hautala N. Prevalence, incidence and epidemiology of childhood uveitis. *Acta Ophthalmol*. 2021;99(2). doi:10.1111/aos.14535
11. De Boer J, Wulffraat N, Rothova A. Visual loss in uveitis of childhood. *British Journal of Ophthalmology*. 2003;87(7). doi:10.1136/bjo.87.7.879
12. Smith JA, Mackensen F, Sen HN, et al. Epidemiology and Course of Disease in Childhood Uveitis. *Ophthalmology*. 2009;116(8). doi:10.1016/j.ophtha.2009.05.002
13. Leal I, Steeples LR, Wong SW, et al. Update on the systemic management of non-infectious uveitis in children and adolescents. *Surv Ophthalmol*. Published online January 2023. doi:10.1016/j.survophthal.2023.01.002
14. Jabs DA, Nussenblatt RB, Rosenbaum JT. Standardization of uveitis nomenclature for reporting clinical data. Results of the First International Workshop. *American journal of ophthalmology*. 2005;140(3):509-516. doi:10.1016/j.ajo.2005.03.057
15. Liu X, Solebo AL, Faes L, et al. Instrument-based Tests for Measuring Anterior Chamber Cells in Uveitis: A Systematic Review. *Ocul Immunol Inflamm*. 2020;28(6):898-907. doi:10.1080/09273948.2019.1640883
16. Kempen JH, Ganesh SK, Sangwan VS, Rathinam SR. Interobserver Agreement in Grading Activity and Site of Inflammation in Eyes of Patients with Uveitis. *Am J Ophthalmol*. 2008;146(6):813-818.e1. doi:10.1016/j.ajo.2008.06.004
17. Oo HH, Mohan M, Song W, et al. Anterior chamber inflammation grading methods: A critical review. *Surv Ophthalmol*. Published online October 2023. doi:10.1016/j.survophthal.2023.10.005
18. Liu X, Solebo AL, Faes L, et al. Instrument-based Tests for Measuring Anterior Chamber Cells in Uveitis: A Systematic Review. *Ocul Immunol Inflamm*. 2020;28(6):898-907. doi:10.1080/09273948.2019.1640883

19. Xie Z, Gadepalli C, Cheetham BMG. REFORMULATION AND GENERALISATION OF THE COHEN AND FLEISS KAPPAS. *LIFE: International Journal of Health and Life-Sciences*. 2017;3(3):1-15. doi:10.20319/ijhls.2017.33.115
20. Glaser DA, Jones D, Carruthers J, et al. Epidemiologic analysis of change in eyelash characteristics with increasing age in a population of healthy women. *Dermatologic surgery : official publication for American Society for Dermatologic Surgery [et al]*. 2014;40(11):1208-1213. doi:10.1097/DSS.0000000000000170
21. Keino H, Aman T, Furuya R, et al. Automated Quantitative Analysis of Anterior Segment Inflammation Using Swept-Source Anterior Segment Optical Coherence Tomography: A Pilot Study. *Diagnostics (Basel, Switzerland)*. 2022;12(11). doi:10.3390/diagnostics12112703
22. Maring M, Saraf SS, Blazes M, et al. Grading Anterior Chamber Inflammation with Anterior Segment Optical Coherence Tomography: An Overview. *Ocular immunology and inflammation*. 2022;30(2):357-363. doi:10.1080/09273948.2022.2028289
23. Sorkhabi MA, Potapenko IO, Ilginis T, Alberti M, Cabrerizo J. Assessment of Anterior Uveitis Through Anterior-Segment Optical Coherence Tomography and Artificial Intelligence-Based Image Analyses. *Translational vision science & technology*. 2022;11(4):7. doi:10.1167/tvst.11.4.7
24. Lin J, Yu L, Weng Q, Zheng X. Retinal image quality assessment for diabetic retinopathy screening: A survey. *Multimedia Tools and Applications*. 2020;79(23):16173-16199. doi:10.1007/s11042-019-07751-6
25. Ledenius K, Svensson E, Stålhammar F, Wiklund LM, Thilander-Klang A. A method to analyse observer disagreement in visual grading studies: Example of assessed image quality in paediatric cerebral multidetector CT images. *British Journal of Radiology*. Published online 2010. doi:10.1259/bjr/26723788
26. Karanasios E, Caglic I, Zawaideh JP, Barrett T. Prostate MRI quality: clinical impact of the PI-QUAL score in prostate cancer diagnostic work-up. *The British Journal of Radiology*. 2022;95(1133). doi:10.1259/bjr.20211372

Table 1. Outcome of images

		Upper third (Slices 1-21)	Middle third (Slices 22-42)	Lower third (Slices 43-64)	Total
Good	Grader 1	114 (17%)	190 (28%)	363 (54%)	667
	Grader 2	121 (19%)	155 (25%)	345 (56%)	621
	Grader 3	128 (21%)	152 (25%)	332 (54%)	612
Limited	Grader 1	432 (28%)	685 (44%)	443 (28%)	1560
	Grader 2	362 (24%)	708 (46%)	461 (30%)	1531
	Grader 3	365 (24%)	712 (46%)	474 (30%)	1551
Poor	Grader 1	236 (87%)	18 (7%)	16 (6%)	270
	Grader 2	318 (83%)	28 (7%)	38 (10%)	384
	Grader 3	322 (83%)	28 (7%)	38 (10%)	388

Outcome of the quality of image (good, limited or poor) as graded by three graders and the level at which the image was taken

Table 2. Outcome of artefact

Presence of lashes	Upper third (Slices 1-21)	Middle third (Slices 22-42)	Lower third (Slices 43-64)	Total
Grader 1	555 (58%)	170 (18%)	233 (24%)	958
Grader 2	568 (56%)	170 (17%)	270 (27%)	1008

Grader 3	569 (56%)	168 (17%)	271 (27%)	1008
<b>Presence of glare</b>	<b>Upper third</b>	<b>Middle third</b>	<b>Lower third</b>	<b>Total</b>
Grader 1	23 (6%)	329 (91%)	10 (3%)	362
Grader 2	16 (7%)	204 (89%)	8 (4%)	228
Grader 3	16 (7%)	212 (90%)	8 (3%)	236
<b>Presence of movement</b>	<b>Upper third</b>	<b>Middle third</b>	<b>Lower third</b>	
Grader 1	67 (66%)	15 (15%)	20 (20%)	102
Grader 2	63 (66%)	15 (16%)	17 (18%)	95
Grader 3	67 (66%)	15 (15%)	20 (20%)	102

Proportion of images with each artefact detected by each grader and the level at which image was taken

**Table 3: Outcome of image cropping**

<b>Crop in Right eye</b>	<b>Right side</b>	<b>Left side</b>	<b>Total</b>
Grader 1	103 (28%)	271 (72%)	374
Grader 2	155 (32%)	326 (68%)	481
Grader 3	156 (32%)	334 (68%)	490
<b>Crop in Left eye</b>	<b>Right side</b>	<b>Left side</b>	<b>Total</b>
Grader 1	787 (97%)	22 (3%)	809

Grader 2	862 (98%)	22 (2%)	884
Grader 3	887 (97%)	30 (3%)	917

Proportion of images that were cropped

Journal Pre-proof