Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Non-ophthalmologist agreement with consultant AS-OCT anterior chamber cell counts in paediatric uveitis: an exploratory masked study

Ali S. Ibrahim

Ali.ibrahim8@nhs.net

University of East Anglia

Ekin Kahraman

University of East Anglia

Sean Zhou

Norfolk and Norwich University Hospital

Narman Puvanachandra

Norfolk and Norwich University Hospital

Saad Khan

Norfolk and Norwich University Hospital

Anas Injarie

Norfolk and Norwich University Hospital

Research Article

Keywords: Anterior segment optical coherence tomography, Paediatric uveitis, Juvenile idiopathic arthritis, Task-shifting, Inter-grader agreement, Intraclass correlation coefficient

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11693405/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. Sean Zhou reports personal fees for educational support from Roche and Bausch & Lomb and is director of Intelligent MedPrep, outside the submitted work. The other authors have no relevant financial or non-financial interests to disclose.

Abstract

Purpose

To determine whether non-ophthalmologists can reliably count anterior chamber cells on anterior segment optical coherence tomography (AS-OCT) images in paediatric uveitis.

Methods

Exploratory masked agreement study using stored routine-care AS-OCT images (113 eye-level episodes, 27 patients, 791 B-scans) at a UK teaching hospital. After one 10-minute teaching session, an optometrist, a medical student and a general nurse independently counted anterior chamber cells, as did a reference consultant paediatric ophthalmologist. A second consultant regraded 75 B-scans. Agreement was summarised with the intraclass correlation coefficient, ICC(A,1), and the mean absolute difference (MAD); confidence intervals accounted for repeated episodes from the same patients.

Results

ICCs versus the reference consultant were 0.922 (95% CI 0.884–0.957), 0.923 (0.904–0.973) and 0.918 (0.884–0.963) for the optometrist, student and nurse, with MADs of 0.75, 0.58 and 0.76 cells. Errors grew with count: at consultant counts of > 0–1 cell, 69–100% of episodes were within ± 1 cell, falling to 7–53% above 5 cells. No grader missed a consultant-positive episode (0/73). On the regraded B-scans, each grader agreed less closely with the second consultant (ICC 0.803–0.844) than the consultants agreed with each other (0.901). Median grading time was 90.0–100.0 s per episode versus 77.5 s for the consultant.

Conclusion

After a 10-minute session, three non-ophthalmologists closely reproduced one consultant's AS-OCT cell counts but agreed less well with a second consultant. Delegated counting looks feasible but needs testing in other centres against a multi-expert reference, and acceptable error and clinical safety are still unknown.

Introduction

Paediatric uveitis accounts for a disproportionate burden of preventable visual morbidity. A UK population-based study reported an incidence of 1.89 per 100 000 children, with 44.3% demonstrating visual loss at uveitis diagnosis and anterior uveitis predominating in 76.9% of cases [1]. Juvenile idiopathic arthritis (JIA)-associated uveitis is the commonest extra-articular manifestation of JIA [2, 3],

yet accounts for only 15% of childhood uveitis overall [1]. Long-term monitoring remains essential to detect subclinical flares and guide treatment escalation.

The Standardization of Uveitis Nomenclature (SUN) grading system is the established framework for quantifying anterior chamber inflammation [4]. However, clinical assessment remains subjective. In the SUN classification project, inter-observer agreement among uveitis experts on uveitic diagnostic categories yielded a mean kappa of only 0.39 [5]. Agreement on slit-lamp anterior chamber cell grading is also limited: between uveitis subspecialists, exact agreement was only fair to moderate (kappa 0.34–0.43), improving substantially only when a one-grade discrepancy was permitted [6]. These data reinforce the need for objective quantification.

Anterior segment optical coherence tomography (AS-OCT) enables more objective, quantitative assessment of anterior chamber inflammatory cells from stored digital images, with inter-grader ICCs of 0.87–0.90 between specialists [7, 8] and AS-OCT-positive findings reported in eyes graded clinically as SUN 0 [9]. However, evidence on whether briefly trained non-ophthalmologists can perform AS-OCT cell counting is limited.

Counting small bright dots against a dark background may depend more on careful looking than on clinical judgement. Automated algorithms have shown correlations above 0.99 with manual counts [7, 10], trained non-physician graders already interpret fundus images in diabetic retinopathy screening [11], and patients can acquire usable fundus images themselves [12], suggesting that well-defined image tasks need not be confined to specialists.

Children with JIA are screened for uveitis as often as every three months, and those with uveitis need frequent monitoring for years [3]. Imaging staff can take AS-OCT scans quickly, but the scans still have to be read by an ophthalmologist, which ties image-based monitoring to specialist availability. If trained non-ophthalmologists could count anterior chamber cells reliably on stored images, consultants could concentrate on children with signs of activity, and monitoring could reach places without a paediatric ophthalmologist on site.

Having previously shown that AS-OCT cell counts correlate with clinical SUN grades in paediatric uveitis [13], we next tested whether the image-interpretation task itself could be delegated. That report proposed the evaluation of non-medical graders as future work [13], and the present study addresses this question using new grading results. To our knowledge, this is the first study to compare AS-OCT anterior chamber cell counting between non-ophthalmologist graders and a consultant ophthalmologist. We selected three graders whose backgrounds span different degrees of eye-specific training: an optometrist, a medical student and a general nurse with no ophthalmic training or prior exposure to OCT images. We asked how closely each grader reproduced a reference consultant's counts, how large the counting errors were, and whether agreement changed when a second consultant supplied the reference.

Methods

Study design and ethics

This single-centre exploratory masked agreement study used a retrospective dataset of stored AS-OCT images acquired during routine care at Norfolk and Norwich University Hospital, a UK secondary care teaching hospital with a specialist paediatric uveitis service. AS-OCT is part of routine monitoring there, alongside clinical SUN grading. All grading, cell counts and analyses reported here are new, generated by a separate masked grading programme conducted on the same routine-care image set as our prior AS-OCT–SUN correlation report [13]. The counts in that report were made by a different grader and are not used here. The grading protocol and analysis plan were pre-specified before grading and applied identically to all three non-ophthalmologist graders. Three analyses were added after the initial analysis: patient-level bootstrap confidence intervals, error summaries by count level, and the comparison of graders with the second consultant.

The study analysed anonymised routine-care imaging data, and individual informed consent was not required. The Research Operations Office at Norfolk and Norwich University Hospital confirmed that formal Research Ethics Committee review was not required, and the NHS Health Research Authority/NIHR decision tool confirmed this classification. The study complied with the Declaration of Helsinki and UK data-protection legislation.

Participants

AS-OCT images were acquired between April 2023 and June 2024. We included 113 imaging episodes from 27 patients with paediatric uveitis (19 JIA-associated, 7 idiopathic, 1 Blau syndrome; mean age 10.7 (SD 3.1) years; 10 male, 17 female). An episode was one eye imaged at one clinic visit. The 113 episodes comprised 51 unique eyes, and patients contributed a median of 2 episodes (range 1–16). Inclusion criteria were paediatric uveitis (age 16 years or under), an AS-OCT volume of at least nine B-scans providing seven analysable central B-scans, and a same-day SUN grade. Of 121 available AS-OCT scans, 113 met these criteria and were included. The remaining eight did not provide seven analysable central B-scans [13]. All consecutive eligible episodes during the study period were included. No formal sample size calculation was performed. The cohort is comparable in scale to specialist–specialist AS-OCT validation studies in paediatric uveitis [8, 14], and the achieved 95% confidence intervals for the primary ICCs were 0.07–0.08 wide, which we considered adequate for a feasibility estimate.

Clinical SUN grades were documented following examination by a single paediatric consultant ophthalmologist and extracted retrospectively from the clinical notes [13]. AS-OCT grading took place between February and May 2026, more than one year after the clinical visits (Table 1).

Table 1
Cohort and imaging characteristics

Characteristic	Value
Participants	
Patients	27
Age, mean (SD), years	10.7 (3.1)
Age, range, years	3–16
Male / female	10 (37.0%) / 17 (63.0%)
JIA-associated uveitis	19 (70.4%)
Idiopathic uveitis	7 (25.9%)
Blau syndrome	1 (3.7%)
Imaging	
Imaging episodes (eye-visits)	113
Left-eye / right-eye episodes	58 / 55
Unique eyes	51
Episodes per patient, median (range)	2 (1–16)
B-scans per grader	791
B-scan counts across four graders	3 164
B-scans regraded by second consultant	75
Clinical SUN grade[a]	
Grade 0	49 (43.4%)
Grade 0.5+	22 (19.5%)
Grade 1+	17 (15.0%)
Grade 2+	15 (13.3%)
Grade 3+	10 (8.8%)
Reference consultant AS-OCT count	
> 0 (active)	73 (64.6%)

[a] Per sequential batch (18–19 episodes), the number of episodes was 6–12 at grade 0, 0–7 at 0.5+, 0–5 at 1+, 0–4 at 2 + and 1–3 at 3+. JIA, juvenile idiopathic arthritis; SUN, Standardization of Uveitis Nomenclature

Characteristic	Value
Participants	
0 (inactive)	40 (35.4%)

[a] Per sequential batch (18–19 episodes), the number of episodes was 6–12 at grade 0, 0–7 at 0.5+, 0–5 at 1+, 0–4 at 2 + and 1–3 at 3+. JIA, juvenile idiopathic arthritis; SUN, Standardization of Uveitis Nomenclature

AS-OCT imaging

AS-OCT imaging was performed using the Spectralis OCT system (Heidelberg Engineering, Heidelberg, Germany) with the anterior segment module (axial resolution 3.9 μm, lateral resolution 5.7 μm) [13]. Each volume scan consisted of 9–48 consecutive B-scans centred on the pupil. The seven B-scans were preselected before grading by a study author (ASI) using a predefined anatomical criterion, as in our previous report [13]: the central (fourth) B-scan was the one with the largest visible pupil diameter, and the three B-scans on either side completed the set. Scans were acquired with the device's standard anterior segment preset, and graders viewed the B-scans in Heidelberg Eye Explorer (HEYEX; Heidelberg Engineering). Image settings, including contrast, brightness and gamma, were not adjusted during grading [13]. The exact B-scan spacing of individual volumes and monitor characteristics were not recorded and could not be recovered.

Graders and training

Four graders assessed all episodes independently. The reference grader was a consultant paediatric ophthalmologist. The three non-ophthalmologist graders were an optometrist, a medical student and a nurse. The optometrist worked in the community and had three years of clinical experience, including general OCT use but no AS-OCT cell counting. The medical student had recently completed their final year and had no OCT or cell-counting experience. The nurse was a general (non-ophthalmic) hospital nurse with four years of clinical experience who had never seen OCT or AS-OCT images. All graders were masked to each other's counts and to clinical SUN grades, and none had access to clinical notes. Each grader viewed the same de-identified B-scans on the same laptop, in sittings supervised and timed by ASI.

Before grading, each non-ophthalmologist received one 10-minute teaching session from a study author (ASI) who had been trained in the counting framework by the reference consultant and who did not take part in grading. The session comprised a standardised verbal explanation and a written rule-based framework (Online Resource 1, Material S2). A countable hyperreflective focus (termed a 'cell' in this paper) was defined as a small, discrete, well-circumscribed hyperreflective dot within the anterior chamber, not contiguous with the corneal endothelium, iris or lens surface. This is an image-based definition, and the foci were not confirmed biologically. Example images used in training were taken from

outside the study dataset. An annotated example B-scan from this cohort, showing examples of counted foci and of signals that were not counted, is provided in Online Resource 1 (Fig. S1). No practice cases or feedback were provided.

Grading protocol

All episodes were randomised before grading and graded by every grader in the same sequence, in sittings of approximately six episodes with no more than one sitting per day. For each episode, graders counted the foci meeting this definition on each of the seven B-scans, and the episode-level count was the mean across the seven B-scans. Grading time was recorded by ASI for 30 episodes (the first five of each successive block of 18–19 episodes), from opening the first B-scan to recording the seventh count.

A second consultant paediatric ophthalmologist, who is also a study author, regraded 75 B-scans from 50 episodes in 20 patients. The B-scans were chosen by random sampling within SUN grades (33, 11, 8, 17 and 6 B-scans at grades 0, 0.5+, 1+, 2 + and 3+). The second consultant was masked to all study counts and clinical SUN grades and used the same written framework.

Statistical analysis

Primary confidence intervals (CIs) for agreement statistics are percentile patient-level bootstrap intervals, which allow for patients contributing several episodes: patients were resampled with replacement, retaining all of their episodes (2000 resamples). These intervals are asymmetric because the sampling distribution of the ICC is skewed. F-based intervals are given in Online Resource 1. Counts are expressed as the mean number of cells per B-scan.

Agreement with the reference consultant was summarised with the intraclass correlation coefficient for absolute agreement of single measures from a two-way random-effects model, ICC(A,1) [15]. We used absolute agreement, not consistency, because a delegated grader's counts would be used as they stand, so systematic over- or undercounting matters as well as ranking. ICCs were interpreted using conventional thresholds (excellent > 0.90, good 0.75–0.90) [16].

Counting error was summarised as the mean absolute difference (MAD) and mean signed difference (bias) between each grader and the reference consultant, with Bland–Altman plots [17]. The horizontal 95% limits of agreement are presented as conventional descriptive limits, because errors varied with count magnitude. MAD and the proportion of episodes within ± 1 cell were therefore also reported by the reference consultant's count (0, > 0–1, > 1–5 and > 5 cells), so that the bands do not depend on the grader being assessed. The ± 1-cell margin is for illustration and is not a validated limit of acceptable error.

To assess whether agreement depended on the reference, each grader's ICC with each consultant was calculated on the second-consultant subsample at B-scan level and compared with the inter-consultant ICC.

Detection was assessed at an any-cell threshold. This is an analytical cut-off, not a clinical one. A false negative was an episode in which the reference consultant (or, separately, the clinical SUN grade) indicated activity and the grader counted no cells. False positives were defined in the same way. Wilson 95% CIs are given for the episode-level false-negative proportion, which assumes independent episodes, and separately for the proportion of patients with any missed episode.

Pre-specified secondary analyses of counting error across the grading sequence, grading-time trends and sensitivity to within-patient clustering are reported in Online Resource 1. Analyses were performed in Python 3.9 using NumPy 2.0, SciPy 1.13 and statsmodels 0.14, with ICCs cross-checked against pingouin 0.5.5. Figures were produced with Matplotlib.

Results

Cohort

The SUN distribution of the 113 episodes was 49 grade 0 (43.4%), 22 grade 0.5+ (19.5%), 17 grade 1+ (15.0%), 15 grade 2+ (13.3%) and 10 grade 3+ (8.8%). The reference consultant counted at least one cell in 73 episodes (64.6%) (Table 1).

How closely did each grader reproduce the reference consultant’s counts?

ICC(A,1) versus the reference consultant was 0.922 (95% CI 0.884–0.957) for the optometrist, 0.923 (0.904–0.973) for the medical student and 0.918 (0.884–0.963) for the nurse (Table 2; Fig. 1). Excluding episodes in which the consultant counted no cells, ICCs were 0.905, 0.909 and 0.900.

How large were the counting errors?

MADs were 0.75 cells (95% CI 0.43–1.00) for the optometrist, 0.58 (0.29–0.87) for the medical student and 0.76 (0.47–1.08) for the nurse. Mean bias was positive for all three (+ 0.42, + 0.15 and + 0.43 cells), and conventional 95% limits of agreement were − 2.46 to + 3.30, − 2.88 to + 3.18 and − 2.54 to + 3.39 cells (Fig. 2). Because errors varied with count magnitude, these horizontal limits describe agreement only approximately.

By consultant count (Table 2), all graders were within ± 1 cell in every episode the consultant scored as zero. At consultant counts of > 0–1 cell, 86%, 100% and 69% of episodes, respectively, were within ± 1 cell; at > 1–5 cells, 79%, 93% and 72%; and above 5 cells, MADs rose to 3.25, 3.06 and 2.36 cells. At SUN grade 0.5+, the nurse’s MAD was 0.85 cells, compared with 0.29 for the optometrist and 0.25 for the medical student (Online Resource 1, Table S4).

Table 2
Agreement with the reference consultant

	Optometrist	Medical student	Nurse
All episodes (n = 113, 27 patients)			
ICC(A,1) (95% CI)	0.922 (0.884–0.957)	0.923 (0.904–0.973)	0.918 (0.884–0.963)
MAD, cells (95% CI)	0.75 (0.43–1.00)	0.58 (0.29–0.87)	0.76 (0.47–1.08)
Bias, cells (95% CI)	+ 0.42 (+ 0.14 to + 0.73)	+ 0.15 (− 0.04 to + 0.39)	+ 0.43 (+ 0.18 to + 0.73)
Conventional 95% limits of agreement, cells	−2.46 to + 3.30	−2.88 to + 3.18	−2.54 to + 3.39
ICC(A,1), consultant-positive episodes (n = 73)	0.905 (0.859–0.948)	0.909 (0.883–0.967)	0.900 (0.858–0.952)
Error by consultant count: MAD; % within ± 1 cell (n)[a]			
0 cells	0.01; 100% (40)	0.01; 100% (40)	0.01; 100% (40)
> 0–1 cell	0.58; 86% (29)	0.24; 100% (29)	0.91; 69% (29)
> 1–5 cells	0.65; 79% (29)	0.41; 93% (29)	0.80; 72% (29)
> 5 cells	3.25; 7% (15)	3.06; 20% (15)	2.36; 53% (15)
Detection at any-cell threshold			
Missed consultant-positive episodes (n = 73)	0 (95% CI 0.0–5.0%)	0 (95% CI 0.0–5.0%)	0 (95% CI 0.0–5.0%)
Patients with a missed episode (n = 18)	0 (95% CI 0.0–17.6%)	0 (95% CI 0.0–17.6%)	0 (95% CI 0.0–17.6%)
Cells counted in consultant-negative episodes (n = 40)	3 (7.5%)	3 (7.5%)	2 (5.0%)
Missed SUN-active episodes (n = 64)[b]	1 (1.6%)	1 (1.6%)	1 (1.6%)
Cells counted in SUN grade 0 episodes (n = 49)[c]	13 (26.5%)	13 (26.5%)	12 (24.5%)

CIs for ICC, MAD and bias are percentile patient-level bootstrap intervals; CIs for missed episodes are Wilson intervals; the episode-level interval assumes independent episodes, and the patient-level row is a separate outcome (any missed episode per patient). [a] Bands defined by the reference consultant's count (mean cells per B-scan); ±1 cell is an illustrative reporting margin. [b] The same single episode for all three graders; the reference consultant also counted zero. [c] In 10 of these episodes the reference consultant

also counted cells. CI, confidence interval; ICC, intraclass correlation coefficient; MAD, mean absolute difference; SUN, Standardization of Uveitis Nomenclature

Did agreement change with a different consultant as reference?

On the 75-B-scan subsample, the two consultants agreed with ICC 0.901 (95% CI 0.794–0.950). The second consultant counted on average 0.31 cells fewer per B-scan, with larger differences at higher grades (Online Resource 1, Table S3). Each grader agreed less closely with the second consultant (ICC 0.803, 0.844 and 0.805 for the optometrist, medical student and nurse) than the consultants agreed with each other, and for all three the 95% CI for the difference excluded zero (Table 3). Against the reference consultant, ICCs on the same B-scans were 0.889, 0.964 and 0.892, close to the inter-consultant value but with wide intervals.

Table 3
Agreement with each consultant on the second-consultant subsample (75 B-scans, 50 episodes, 20 patients)

Second vs reference consultant	ICC(A,1) (95% CI)	Difference from inter-consultant ICC (95% CI)
	0.901 (0.794–0.950)	**—**
Versus reference consultant		
Optometrist	0.889 (0.846–0.940)	−0.011 (− 0.087 to + 0.119)
Medical student	0.964 (0.938–0.988)	+ 0.064 (− 0.001 to + 0.185)
Nurse	0.892 (0.827–0.934)	−0.008 (− 0.064 to + 0.080)
Versus second consultant		
Optometrist	0.803 (0.731–0.841)	−0.098 (− 0.147 to − 0.030)
Medical student	0.844 (0.697–0.908)	−0.057 (− 0.122 to − 0.024)
Nurse	0.805 (0.695–0.871)	−0.095 (− 0.142 to − 0.057)

B-scan-level counts; percentile patient-level bootstrap 95% CIs. Post hoc analysis. ICC, intraclass correlation coefficient

Detection at the any-cell threshold

No grader recorded a false negative against the reference consultant: 0/73 episodes (Wilson 95% CI 0.0–5.0%, assuming independent episodes), and no patient with any missed episode (0/18; 95% CI 0.0–17.6%). Against clinical SUN grading, each grader missed the same single SUN 0.5 + episode, in which the reference consultant also counted no cells. Graders counted cells in 3, 3 and 2 of the 40 episodes the consultant scored as zero, each with one or two cells across seven B-scans (Table 2).

Grading time

Median time per episode was 77.5 s for the consultant, 90.5 s for the optometrist, 100.0 s for the medical student and 90.0 s for the nurse. Secondary analyses found no clear trend in counting error over the grading sequence for any grader (adjusted change per 10 episodes − 0.049, − 0.015 and − 0.008 cells, all 95% CIs including zero), and all graders became faster (Online Resource 1, Table S5).

Discussion

After a single 10-minute session, an optometrist, a medical student and a general nurse reproduced a consultant's AS-OCT cell counts closely (ICC 0.918–0.923). Their counts were less close to those of a second consultant. In other words, the graders' counts were closer to the reference consultant's than to the second consultant's, so agreement with one consultant should not be read as agreement with consultants in general. These data cannot tell us why the gap arose. Without baseline testing, we cannot say what the training itself achieved. The graders were taught by an author who had learned the method from the reference consultant, and the second consultant tended to count fewer cells in busy eyes; either may have contributed. The comparison also rests on 75 single B-scans. It shows that the choice of reference matters, not which consultant is nearer the truth. Any service using delegated counting would need training and quality checks based on more than one expert.

Performance did not follow the graders' level of ophthalmic experience. The medical student had the smallest overall errors, and the nurse, who had not seen an OCT image before the study, did about as well as the optometrist. This fits with cell counting being mainly a visual task with clear rules, which is also why automated counting performs well [7, 10, 18] and why trained non-physicians can grade retinal photographs [11]. The nurse did count more generously in quiet eyes. When the consultant counted more than 0 and up to 1 cell per B-scan, the nurse was within one cell in 69% of episodes (MAD 0.91 cells), against 100% for the medical student. Low-grade activity is exactly what monitoring is meant to pick up, so this is an individual tendency that calibration against reference images would need to correct. With one grader from each background, none of these differences can be put down to profession.

The ICCs are high partly because the counts covered a wide range, from many eyes with no cells to a few with dozens. A high ICC can sit alongside errors that matter clinically, and the absolute differences and count-level results give a better sense of how far out a grader might be. Average bias was half a cell or less. Individual discrepancies were larger in inflamed eyes (MAD 2.36–3.25 cells above 5 cells per B-scan), and together with the nurse's overcounting at low counts, they argue against treating these

graders' counts as interchangeable with the consultant's for longitudinal monitoring. What level of error is acceptable has not yet been defined.

No grader missed an episode in which the consultant saw cells. That is agreement on detection, not proof of safety: with 18 patients the upper confidence limit is 18%, which leaves room for a clinically important miss rate. The one SUN-active eye that every grader scored as clear was also scored as clear by the consultant, which is consistent with an imaging or sampling limitation, such as looking at only seven B-scans. Cells counted in eyes that looked quiet at the slit lamp [9] need follow-up to show whether they matter. In any service, a positive scan should lead to clinical review, not to a change in treatment.

Hypothetically, imaging staff could take scans, trained graders could count them, and scans with cells, or a rise since the last visit, could be referred to a consultant. Whether this would save consultant time is unknown: the graders here were slower than the consultant (median 90.0–100.0 s against 77.5 s per episode), and acquisition, frame selection, supervision and quality assurance were not evaluated.

The study has clear limits. It was done in one centre, on one device, with images selected because they were analysable. There were no SUN 4 + eyes, and only 8.8% of episodes were graded 3+. With one grader from each background, it shows what three individuals achieved, not what a workforce would achieve, and it does not show that 10 minutes of training is enough in general. We did not test ophthalmic nurses, who are perhaps the most likely people to take on this work. A single consultant provided the reference, both consultants are authors, and the second-consultant comparison was added after the main analysis and covers only a subset of B-scans. Timing was recorded for 30 episodes, and the graders knew they were being timed. Some hyperreflective foci may be pigment or debris [19]. Finally, the graders were given pre-selected B-scans, so we have not tested whether non-ophthalmologists can take scans or choose the right frames themselves.

The next step is a multicentre study with the staff who would actually do the work, several graders from each group, a reference agreed by more than one expert, longitudinal outcomes, and a comparison with automated counting [20].

Conclusion

After a 10-minute training session, an optometrist, a medical student and a general nurse closely reproduced one consultant's counts of hyperreflective foci on pre-selected AS-OCT images, but agreed less well with a second consultant. Delegated counting looks feasible. It now needs testing in other centres against a reference set by more than one expert, and it is not yet known how much counting error is acceptable or whether the approach is safe in a clinical pathway.

Declarations

Funding No funding was received for conducting this study.

Competing interests Sean Zhou reports personal fees for educational support from Roche and Bausch & Lomb and is director of Intelligent MedPrep, outside the submitted work. The other authors have no relevant financial or non-financial interests to disclose.

Author contributions Ali S. Ibrahim conceived and designed the study, delivered the training to the non-specialist graders, wrote the manuscript and supervised the project. Ekin Kahraman performed the statistical analysis and contributed to writing the manuscript. Sean Zhou contributed to planning the study, data curation and analytical support, and reviewed the manuscript. Narman Puvanachandra performed the second-consultant regrading of the 75-B-scan subsample and reviewed the manuscript. Saad Khan contributed to interpretation of the findings and critically revised the manuscript for important intellectual content. Anas Injarie performed reference grading, trained Ali S. Ibrahim in the counting framework, established and led the paediatric uveitis AS-OCT imaging programme from which the study cohort was derived, contributed to study conception and design, provided clinical oversight of image acquisition and data collection, contributed to interpretation of the findings and critically revised the manuscript. All authors read and approved the final manuscript.

Ethics approval This retrospective study of anonymised routine-care imaging data was performed in line with the principles of the Declaration of Helsinki. The Research Operations Office at Norfolk and Norwich University Hospital confirmed that formal Research Ethics Committee review was not required, and the NHS Health Research Authority/NIHR decision tool confirmed this classification.

Consent to participate Individual informed consent was not required for this analysis of anonymised data collected during routine clinical care.

Consent to publish No identifiable patient information is included. The example AS-OCT B-scan in Online Resource 1 is de-identified.

Data availability The datasets analysed during this study are available from the corresponding author on reasonable request, subject to institutional approval and patient confidentiality requirements.

References

1. 1. Solebo AL, Kellett S, McLoone E et al (2025) Incidence, sociodemographic and presenting clinical features of childhood non-infectious uveitis: findings from the UK national inception cohort study. Br J Ophthalmol 109:955–961. https://doi.org/10.1136/bjo-2024-326674
2. 2. Clarke SLN, Sen ES, Ramanan AV (2016) Juvenile idiopathic arthritis-associated uveitis. Pediatr Rheumatol Online J 14:27. https://doi.org/10.1186/s12969-016-0088-2
3. 3. Angeles-Han ST, Ringold S, Beukelman T et al (2019) 2019 American College of Rheumatology/Arthritis Foundation guideline for the screening, monitoring, and treatment of juvenile idiopathic arthritis-associated uveitis. Arthritis Care Res (Hoboken) 71:703–716. https://doi.org/10.1002/acr.23871

4. 4. Jabs DA, Nussenblatt RB, Rosenbaum JT, Standardization of Uveitis Nomenclature (SUN) Working Group (2005) Standardization of uveitis nomenclature for reporting clinical data. Results of the First International Workshop. Am J Ophthalmol 140:509–516. https://doi.org/10.1016/j.ajo.2005.03.057
5. 5. Jabs DA, Dick AD, Doucette JT et al (2018) Interobserver agreement among uveitis experts on uveitic diagnoses: the Standardization of Uveitis Nomenclature experience. Am J Ophthalmol 186:19–24. https://doi.org/10.1016/j.ajo.2017.10.028
6. 6. Kempen JH, Ganesh SK, Sangwan VS, Rathinam SR (2008) Interobserver agreement in grading activity and site of inflammation in eyes of patients with uveitis. Am J Ophthalmol 146:813–818.e1. https://doi.org/10.1016/j.ajo.2008.06.004
7. 7. Sharma S, Lowder CY, Vasanji A et al (2015) Automated analysis of anterior chamber inflammation by spectral-domain optical coherence tomography. Ophthalmology 122:1464–1470. https://doi.org/10.1016/j.ophtha.2015.02.032
8. 8. Tsui E, Chen JL, Jackson NJ et al (2022) Quantification of anterior chamber cells in children with uveitis using anterior segment optical coherence tomography. Am J Ophthalmol 241:254–261. https://doi.org/10.1016/j.ajo.2022.05.012
9. 9. Pillar S, Kadomoto S, Chen K et al (2025) Automated quantification of anterior chamber cells using swept-source anterior segment optical coherence tomography. J Ophthalmic Inflamm Infect 15:3. https://doi.org/10.1186/s12348-025-00456-y
10. 10. Kang TS, Lee Y, Lee S et al (2021) Development of fully automated anterior chamber cell analysis based on image software. Sci Rep 11:10670. https://doi.org/10.1038/s41598-021-89794-0
11. 11. Bhargava M, Cheung CY, Sabanayagam C et al (2012) Accuracy of diabetic retinopathy screening by trained non-physician graders using non-mydriatic fundus camera. Singapore Med J 53:715–719
12. 12. Kumari S, Venkatesh P, Tandon N et al (2022) Selfie fundus imaging for diabetic retinopathy screening. Eye (Lond) 36:1988–1993. https://doi.org/10.1038/s41433-021-01804-7
13. 13. Ibrahim AS, Tam T, Zhou S et al (2026) Correlation between SUN scale grades and AS-OCT cell counts in pediatric uveitis: a retrospective central 7-frame analysis. Clin Ophthalmol 20:575712. https://doi.org/10.2147/OPTH.S575712
14. 14. Uthayananthan P, Tanwar N, Rahi JS et al (2025) Imaging-based detection of anterior chamber inflammation: a comparative diagnostic accuracy study. Am J Ophthalmol 270:131–139. https://doi.org/10.1016/j.ajo.2024.07.018
15. 15. McGraw KO, Wong SP (1996) Forming inferences about some intraclass correlation coefficients. Psychol Methods 1:30–46. https://doi.org/10.1037/1082-989X.1.1.30
16. 16. Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15:155–163. https://doi.org/10.1016/j.jcm.2016.02.012
17. 17. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327:307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

18. 18. Sorkhabi MA, Potapenko IO, Ilginis T et al (2022) Assessment of anterior uveitis through anterior-segment optical coherence tomography and artificial intelligence-based image analyses. Transl Vis Sci Technol 11:7. https://doi.org/10.1167/tvst.11.4.7
19. 19. Bellchambers A, Shu R, Chu CJ et al (2025) Differentiation of anterior chamber pigment and inflammatory cells using swept-source optical coherence tomography: a cross-sectional study. Eye (Lond) 39:1744–1750. https://doi.org/10.1038/s41433-025-03697-2
20. 20. Patel A, Acharya NR, Solebo AL (2026) Novel advances in juvenile idiopathic arthritis associated uveitis. Arthritis Res Ther 28:62. https://doi.org/10.1186/s13075-026-03758-1

Figures

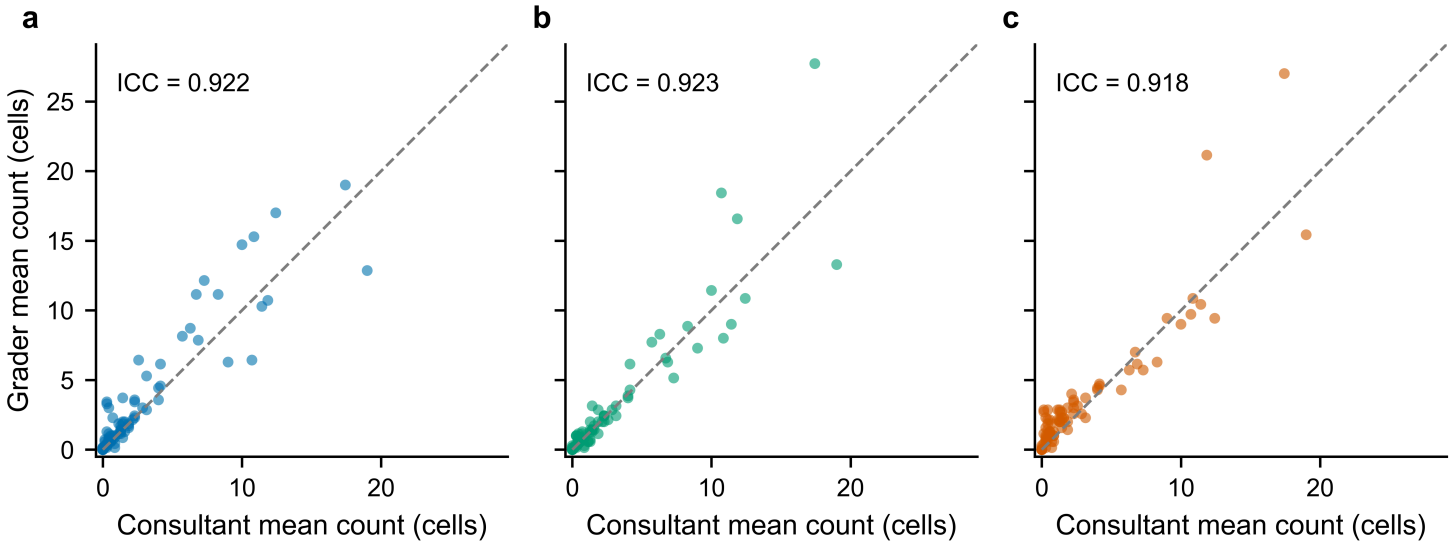

Figure 1

Episode-level mean cell counts of (a) the optometrist, (b) the medical student and (c) the nurse plotted against the reference consultant for 113 episodes, with the ICC(A,1) for each grader. Dashed line, line of identity

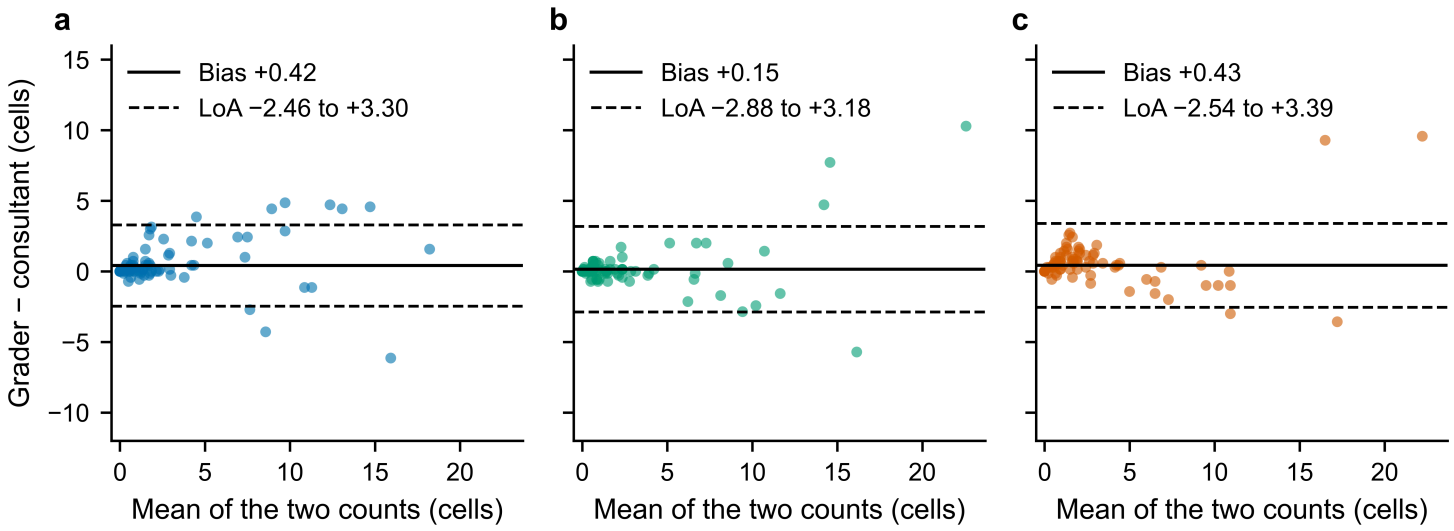

Figure 2

Bland–Altman plots for (a) the optometrist, (b) the medical student and (c) the nurse versus the reference consultant (113 episodes). Solid line, mean bias; dashed lines, conventional descriptive 95% limits of agreement (LoA). Differences widen as counts increase, so the horizontal limits describe agreement only approximately

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf