Scientific Research Publishing

Open Journal of Ophthalmology, 2026, 16(3), 254-266
https://www.scirp.org/journal/ojoph
ISSN Online: 2165-7416
ISSN Print: 2165-7408

Pre-Biopsy Diagnostic Accuracy for Orbital Lymphoma versus Inflammatory Orbital Disease: A 10-Year Retrospective Study at a UK Tertiary Centre

Kit May Chow[1*], Bernard Chang[1], George Kalantzis[1], Sreedhar Jyothi[1], Nabil El-Hindy[1], Nithikka Senthilkumar[2], Bahtigul Arzikulova[2], Gabriella Guevara[1]

[1]Department of Ophthalmology, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom
[2]School of Medicine, University of Leeds, Leeds, United Kingdom
Email: *catherinekmchow@gmail.com

How to cite this paper: Chow, K.M., Chang, B., Kalantzis, G., Jyothi, S., El-Hindy, N., Senthilkumar, N., Arzikulova, B. and Guevara, G. (2026) Pre-Biopsy Diagnostic Accuracy for Orbital Lymphoma versus Inflammatory Orbital Disease: A 10-Year Retrospective Study at a UK Tertiary Centre. *Open Journal of Ophthalmology*, **16**, 254-266.
https://doi.org/10.4236/ojoph.2026.163022

Received: June 24, 2026
Accepted: July 31, 2026
Published: August 3, 2026

Abstract

Purpose: To determine the diagnostic accuracy of pre-biopsy clinical and radiological assessment in differentiating orbital lymphoma from inflammatory orbital disease, and to identify clinical, anatomical, and pathological predictors of diagnostic failure. **Methods:** A single-center retrospective diagnostic accuracy study (STARD 2015-compliant) of 88 consecutive orbital biopsies was performed at Leeds Teaching Hospitals NHS Trust between January 2014 and December 2024. Cases were classified as lymphoma (n = 36) or inflammatory disease (n = 52) based on histopathology. Index tests were pre-biopsy clinical and radiological impressions. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and likelihood ratios were calculated with 95% confidence intervals. Age was evaluated as a continuous discriminator using receiver operating characteristic (ROC) analysis. Presenting features, anatomical compartment, and symptom duration were compared between groups. **Results:** Clinical assessment achieved sensitivity of 55.6% (95% CI 39.6% - 70.5%) and specificity of 90.4% (95% CI 79.4% - 95.8%). Radiological assessment achieved a sensitivity of 61.1% (95% CI 44.9% - 75.2%) and a specificity of 88.5% (95% CI 77.0% - 94.6%). Age was a strong discriminator (area under the ROC curve 0.874, 95% CI 0.793 - 0.954; $p < 0.001$); an optimal threshold of ≥55 years yielded sensitivity of 94.4% and NPV of 95.0%. Twelve lymphoma cases (33.3%) were missed by both assessors (false negative [FN]-Both). Extraconal lymphoma had the highest FN-Both rate (44.4%). Proptosis was significantly more prevalent in lymphoma (41.7% vs 17.3%; $p = 0.015$). Lymphoma cases had a shorter median symptom duration (5 vs 6

months; Mann-Whitney p = 0.008), reflecting the chronic relapsing course of IgG4-related disease and sarcoidosis in the inflammatory group. Four systematic failure patterns were identified. **Conclusions:** Pre-biopsy assessment reliably raises suspicion of orbital lymphoma when positive but cannot safely exclude it when negative. Age over 55, extraconal location, and proptosis are key risk factors. Orbital biopsy remains indispensable regardless of pre-test impression.

Keywords

Orbital Lymphoma, Orbital Biopsy, Diagnostic Accuracy, Orbital Inflammatory Disease, Sensitivity, Specificity

1. Introduction

Orbital biopsy remains the definitive diagnostic procedure for orbital lesions of uncertain aetiology, with histopathology constituting the reference standard against which all clinical and radiological assessments must be judged [1]. Among the most clinically challenging orbital diagnoses is the differentiation of orbital lymphoma from inflammatory orbital disease. Both entities may present with similar signs: proptosis, periorbital swelling, palpable mass, and diplopia, and may share overlapping radiological characteristics, particularly when lesions arise in the extraconal compartment or involve the lacrimal gland [2] [3].

This diagnostic uncertainty carries significant clinical consequences. Orbital lymphoma demands systemic haematological workup, staging, and prompt oncological treatment, whereas inflammatory orbital disease, encompassing idiopathic orbital inflammation (orbital pseudotumour), IgG4-related orbital disease, and sarcoidosis, is managed with immunosuppression. Delayed diagnosis of lymphoma following an assumption of inflammation risks deferred systemic treatment and potential disease progression. Conversely, over-investigation of benign inflammatory conditions may expose patients to unnecessary staging procedures, radiation, and associated anxiety.

Prior studies of overall orbital biopsy accuracy have consistently reported correct pre-biopsy diagnoses in fewer than 50% of all orbital lesion cases, with inflammatory and haematological lesions identified as the most diagnostically challenging subgroup [1] [2] [4]-[6]. Koukkoulli *et al.* at our institution previously demonstrated that the greatest challenge in orbital biopsy practice lay precisely in differentiating inflammatory from haematological lesions, which represented half of their cases [2]. However, formal diagnostic accuracy metrics such as sensitivity, specificity, and likelihood ratios specifically for the lymphoma versus inflammation sub-problem have not been rigorously quantified in a dedicated study, and the anatomical and clinical predictors of diagnostic failure in this subgroup remain poorly characterised.

We therefore conducted a dedicated retrospective diagnostic accuracy study of

88 consecutive orbital biopsies over a 10-year period, focusing specifically on cases where lymphoma or inflammatory disease was the principal diagnostic consideration. We aimed to: 1) quantify the sensitivity, specificity, and predictive values of pre-biopsy clinical and radiological impression; 2) evaluate age as a continuous diagnostic discriminator using ROC analysis; 3) characterize the impact of anatomical location on diagnostic failure; and 4) identify clinical features and systematic failure patterns that may inform future orbital biopsy practice.

2. Materials and Methods

2.1. Study Design and Setting

This was a single-centre retrospective diagnostic accuracy study, reported in compliance with the STARD 2015 guidelines [7]. The study was performed at the Department of Ophthalmology, Leeds Teaching Hospitals NHS Trust. Ethical approval was obtained under institutional clinical audit governance. As a retrospective analysis of existing de-identified clinical data, individual patient consent was not required.

2.2. Patient Selection

All patients who underwent surgical orbital biopsy between January 2014 and December 2024 were identified through institutional surgical records. Inclusion criteria were: 1) orbital biopsy performed for a lesion in which lymphoma or inflammatory orbital disease featured in the pre-biopsy differential diagnosis; and 2) histopathological diagnosis available as the reference standard. Eighty-eight consecutive biopsy encounters in 80 unique patients met the inclusion criteria. Six patients underwent more than one biopsy during the study period; each encounter was retained as an independent study entry, as each involved a distinct pre-biopsy clinical and radiological assessment.

2.3. Index Tests and Reference Standard

Two independent index tests were evaluated for each case: 1) the pre-biopsy clinical impression, as documented in clinical records; and 2) the pre-biopsy radiological impression, as documented in the formal radiology report. Each was classified as "lymphoma suspected" (Yes) or "lymphoma not suspected" (No/Not documented/Not performed). Cases in which no imaging was performed or no radiological impression was documented (n = 19) were coded as "lymphoma not suspected - radiology," consistent with intent-to-diagnose analysis and reflecting real-world clinical conditions. A case was classified as "lymphoma suspected" only if lymphoma was explicitly named as a diagnosis or differential diagnosis in the clinical note or radiology report; non-specific entries (e.g., "rule out malignancy" or purely anatomical descriptions without diagnostic impression) were coded as "lymphoma not suspected." Where a report listed multiple differential diagnoses, including lymphoma, the case was coded as "lymphoma suspected." This coding rule was applied independently to clinical and radiological index tests.

The reference standard was histopathological diagnosis from surgical orbital biopsy. Lymphoma diagnoses were classified according to the 2016 World Health Organization classification of lymphoid neoplasms [8]. Where a patient underwent more than one biopsy, the final confirmed histological diagnosis was used as the reference standard for all of that patient's biopsy entries. In all six patients with repeat biopsies, the histopathological diagnosis was consistent across encounters (*i.e.*, each biopsy yielded the same disease category), so there is no difference between using each encounter's own biopsy result versus the final diagnosis as the reference standard for this cohort.

2.4. Statistical Analysis

Sensitivity, specificity, PPV, NPV, positive likelihood ratio (LR+), negative likelihood ratio (LR−), and Youden's Index were calculated from 2 × 2 contingency tables; 95% confidence intervals were calculated for all proportions. ROC analysis with trapezoidal AUC and Hanley-McNeil standard error was performed for age as a continuous discriminator; the optimal threshold was identified by Youden's Index. Between-group differences in age and symptom duration were assessed by the Mann-Whitney U test; between-group differences in categorical presenting features were assessed by Fisher's exact test. All analyses were performed using SPSS version 26 and Microsoft Excel. Two-tailed tests were applied throughout; statistical significance was defined at $p < 0.05$.

3. Results

3.1. Study Population

Eighty-eight consecutive orbital biopsies in 80 unique patients were included: 36 biopsy events resulted in a diagnosis of orbital lymphoma (40.9%), and 52 resulted in inflammatory orbital disease (59.1%). Six patients (7.5%) underwent more than one biopsy during the study period; each biopsy encounter was included as a separate study entry, as each represents an independent pre-biopsy assessment event. Among patients with repeat biopsies, the indication was disease recurrence in two patients with orbital lymphoma and relapsing or progressive inflammatory orbital disease in four patients (comprising IgG4-related orbital disease, Behçet's disease-associated orbital inflammation, and recurrent non-specific orbital inflammation). Among the lymphoma group, the most common histological subtypes were extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (EMZL/MALT; n = 20, 55.6%), followed by follicular lymphoma (n = 6, 16.7%) and diffuse large B-cell lymphoma (DLBCL; n = 4, 11.1%). The inflammatory group comprised non-specific orbital inflammation (n = 22, 42.3%), sarcoidosis (n = 7, 13.5%), IgG4-related orbital disease (n = 4, 7.7%), and other granulomatous or chronic inflammatory diagnoses. Nineteen biopsy encounters (21.6%) had no formal diagnostic conclusion documented in the pre-biopsy radiological record; all were coded as "lymphoma not suspected - radiology" in the intent-to-diagnose analysis. The largest sub-category comprised 14 cases in which imaging

was performed but the radiology report was purely descriptive, recording dimensions or anatomical location without offering a diagnostic impression. A further two cases had imaging performed at referring external centres with no locally available radiology report. One case had no imaging performed because the orbit had been previously enucleated (prior retinoblastoma), with the lesion being directly accessible for biopsy. The remaining two cases were taken to biopsy on the basis of established clinical indicators without local imaging: one with a known diagnosis of systemic lymphoma involving the orbit, and one with known abnormal paraprotein levels at the time of presentation. Lesions were anteriorly located in both cases.

3.2. Diagnostic Accuracy of Pre-Biopsy Assessment

For clinical impression, the 2 × 2 contingency table yielded: TP = 20, FP = 5, FN = 16, TN = 47. For radiological assessment: TP = 22, FP = 6, FN = 14, TN = 46. Complete diagnostic accuracy metrics with 95% CIs are presented in Table 1.

Clinical assessment demonstrated a sensitivity of 55.6% (95% CI 39.6% - 70.5%) and a specificity of 90.4% (95% CI 79.4% - 95.8%), with LR+ 5.78 and LR− 0.49. Radiological assessment showed comparable performance: sensitivity 61.1% (95% CI 44.9% - 75.2%), specificity 88.5% (95% CI 77.0% - 94.6%), LR+ 5.30, LR− 0.44. Both assessments thus perform well as rule-in tests (LR+ > 5) but perform poorly as rule-out tests (LR− > 0.1). Positive predictive values were 80.0% and 78.6%, respectively, reflecting the near-equal prevalence of lymphoma and inflammation in this biopsy cohort.

Table 1. Diagnostic accuracy metrics for pre-biopsy clinical and radiological assessment (n = 88).

Metric	Clinical Impression	95% CI	Radiological Assessment	95% CI
True Positive	20	—	22	—
False Positive	5	—	6	—
False Negative	16	—	14	—
True Negative	47	—	46	—
Sensitivity	55.6%	39.6% - 70.5%	61.1%	44.9% - 75.2%
Specificity	90.4%	79.4% - 95.8%	88.5%	77.0% - 94.6%
Positive Predictive Value	80.0%	60.9% - 91.1%	78.6%	60.5% - 89.8%
Negative Predictive Value	74.6%	62.7% - 83.7%	76.7%	64.6% - 85.6%
Positive Likelihood Ratio (LR+)	5.78	—	5.30	—
Negative Likelihood Ratio (LR−)	0.49	—	0.44	—
Youden's Index	0.459	—	0.496	—

Abbreviations: CI, confidence interval; LR, likelihood ratio. 95% confidence intervals were applied to all proportions. Nineteen cases had no formal diagnostic conclusion in the pre-biopsy radiological record (14 descriptive-only reports, 2 external referrals without local reports, 3 no imaging performed); all were coded as "lymphoma not suspected - radiology" in the intent-to-diagnose analysis. In a secondary per-protocol analysis excluding these 19 cases (n = 69), radiological sensitivity was 61.1% (95% CI 44.9% - 75.2%), specificity 88.7% (95% CI 76.0% - 95.2%), LR+ 5.41, LR− 0.44 — consistent with the primary intent-to-diagnose results, confirming that the inclusion of non-documented cases did not materially affect the radiological accuracy estimates.

3.3. Age as a Diagnostic Discriminator

Lymphoma cases were significantly older than inflammatory cases (median age 74 years, IQR 71 - 80 vs 48 years, IQR 36 - 57; Mann-Whitney U = 243, z = 5.88, $p < 0.001$). ROC analysis demonstrated an AUC of 0.874 (95% CI 0.793 - 0.954) for age as a discriminator for lymphoma. The optimal Youden threshold was age ≥ 55 years (Youden's J = 0.675), yielding sensitivity 94.4%, specificity 73.1%, PPV 70.8%, and NPV 95.0%. Full results are presented in Table 2.

Table 2. Age as a continuous discriminator for orbital lymphoma—mann-whitney U test and ROC analysis (n = 88).

Parameter	Value
Lymphoma - median age (IQR), years	74 (71 - 80)
Inflammatory disease—median age (IQR), years	48 (36 - 57)
Mann-Whitney U statistic	243.0
Mann-Whitney z-score	5.88
p-value	<0.001
ROC Area Under the Curve (AUC)	0.874
AUC 95% CI (Hanley-McNeil)	0.793 - 0.954
Optimal age threshold (Youden's index)	≥55 years
Sensitivity at threshold	94.4%
Specificity at threshold	73.1%
PPV at threshold	70.8%
NPV at threshold	95.0%
Youden's J at threshold	0.675

Abbreviations: IQR, interquartile range; ROC, receiver operating characteristic; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value. AUC calculated using the trapezoidal method with Hanley-McNeil standard error. The optimal threshold was identified by maximising Youden's Index.

3.4. Anatomical Location and Diagnostic Failure

Location data were available for all cases. Among lymphoma cases, diffuse orbital involvement was the most common pattern (n = 16, 44.4%), followed by lacrimal gland (n = 10, 27.8%) and extraconal (n = 9, 25.0%); intraconal involvement occurred in only one case. In the inflammatory group, lacrimal gland involvement predominated (n = 27, 54.0%).

The FN-Both rate varied substantially by anatomical compartment (Table 3). Extraconal lymphoma had the highest missed-diagnosis rate at 44.4% (4/9), followed by the lacrimal gland at 30.0% (3/10), and diffuse disease at 12.5% (2/16). The single intraconal case was also missed, though this represents insufficient numbers for interpretation. These findings indicate that extraconal location is a specific risk factor for pre-biopsy missed lymphoma diagnosis.

Table 3. Anatomical compartment of orbital lesion and fn-both rate in the lymphoma group (n = 36).

Compartment	Lymphoma Cases, n (%)	Inflammation Cases, n (%)	FN-Both, n (%)	Notes
Lacrimal gland	10 (27.8%)	26 (50.0%)	3 (30.0%)	Substantial miss rate
Extraconal	9 (25.0%)	9 (17.3%)	4 (44.4%)	Highest miss rate; lymphoma under-recognized
Diffuse	16 (44.4%)	16 (30.8%)	2 (12.5%)	Most common lymphoma pattern; lowest miss rate
Intraconal	1 (2.8%)	1 (1.9%)	1 (100%)	N = 1; interpret with caution

FN-Both: missed by both clinical and radiological assessors.

3.5. Presenting Features and Symptom Duration

Presenting feature and symptom duration data are summarised in Table 4. Proptosis was significantly more prevalent in the lymphoma group (41.7%, 15/36) than in the inflammatory group (17.3%, 9/52; Fisher's exact p = 0.015). Pain trended toward greater frequency in the inflammatory group (19.2% vs 5.6%) but did not reach statistical significance (p = 0.113). Palpable mass was the most common presenting feature in both groups and did not discriminate between them (75.0% vs 69.2%; p = 0.635). Diplopia was present in 25.0% of lymphoma and 15.4% of inflammatory cases (p = 0.280).

Table 4. Presenting features and symptom duration by diagnosis group.

Feature	Lymphoma n/N (%)	Inflammation n/N (%)	p-value[a]	Comment
Proptosis	15/36 (41.7%)	9/52 (17.3%)	0.015	Significant; favours lymphoma
Palpable mass	27/36 (75.0%)	36/52 (69.2%)	0.635	Not discriminating
Diplopia	9/36 (25.0%)	8/52 (15.4%)	0.280	Not significant
Pain	2/36 (5.6%)	10/52 (19.2%)	0.113	Trend; favours inflammation
Median symptom duration (IQR), months	5 (2 - 8)	6 (4 - 14)	0.008[b]	Shorter duration in lymphoma
Prior steroid treatment	3/36 (8.3%)	10/52 (19.2%)	0.219	Not significant

[a] Fisher's exact test (two-tailed). [b] Mann-Whitney U test (two-tailed). IQR, interquartile range.

Symptom duration was defined as the interval from first patient-reported symptom onset (as recorded in the clinical note at the time of initial presentation) to the date of orbital biopsy. For patients with relapsing disease who underwent repeat biopsy, duration was calculated from the onset of the most recent relapse to the date of the relevant biopsy encounter. Median symptom duration to biopsy was shorter in the lymphoma group (5 months, IQR 2 - 8 months) than in the inflammatory group (6 months, IQR 4 - 14 months; Mann-Whitney p = 0.008). Notably, the mean duration in the inflammatory group was substantially higher (17.2 months vs 5.5 months), reflecting the relapsing clinical course of IgG4-re-

lated disease and sarcoidosis. Prior steroid treatment (post imaging) was documented in 3 of 36 lymphoma cases (8.3%) and 10 of 52 inflammatory cases (19.2%; p = 0.219).

3.6. Failure Analysis

Twelve of 36 lymphoma cases (33.3%) were missed by both clinical and radiological assessors simultaneously. A systematic review of these 12 cases identified four recurring failure patterns:

1) Descriptive or non-diagnostic impression (n = 5): In all five cases, imaging was available, but neither assessor offered a diagnostic conclusion. Radiology reports recorded only anatomical findings—mass dimensions, lacrimal gland enlargement, or enhancement pattern—without a differential diagnosis. Clinical impressions were similarly non-specific, documenting anatomy or a non-lymphoma inflammatory label without raising lymphoma as a possibility.

2) Competing pathology misattribution (n = 2): An alternative diagnosis was actively proposed by one or both assessors—orbital varices in one case, pulmonary carcinoma metastasis in another—displacing lymphoma from the differential entirely.

3) Prior-diagnosis anchoring (n = 2): A pre-existing diagnosis (idiopathic orbital inflammation; prior biopsy-confirmed inflammatory disease) led assessors to interpret new or progressive orbital disease as recurrence, without reconsidering lymphoma.

4) No pre-biopsy imaging or report available (n = 3): No diagnostic impression was documented by either assessor. In several cases, no imaging had been performed prior to biopsy as the lesions were anterior/extraconal, or a report from an external centre was not available for review. One case had known abnormal serum paraproteins pre-biopsy, but lymphoma was not raised as a diagnostic consideration.

3.7. Complication Rates

Intraoperative or perioperative complications were recorded in 4 of 36 lymphoma biopsies (11.1%) and 2 of 52 inflammatory biopsies (3.8%). Complications in the lymphoma group included conjunctival chemosis with resultant 1 mm lagophthalmos (n = 1), restricted extraocular movement (n = 1), supraorbital numbness (n = 1), and chronic disc swelling (n = 1). In the inflammatory group, complications included transient diplopia (n = 1) and a post-operative orbital haematoma (n = 1). There was no operative mortality. The overall cohort complication rate was 6.8% (6/88).

4. Discussion

This study provides the first dedicated quantitative assessment of pre-biopsy clinical and radiological diagnostic accuracy for the lymphoma versus inflammatory disease sub-problem in a UK tertiary hospital. The central finding is that both

assessment modalities function well as rule-in tools, with LR+ values exceeding 5; a positive impression represents a clinically meaningful shift in post-test probability, but they function poorly as rule-out tools, with LR− values of approximately 0.4 indicating an insufficient reduction in post-test probability when the impression is negative. This pattern mirrors established findings across the diagnostic accuracy literature, where subjective expert impression outperforms its negative performance [1].

A sensitivity of 55.6% for clinical impression and 61.1% for radiological assessment means that more than one-third of lymphoma cases were not suspected pre-biopsy by each modality individually. The FN-Both rate of 33.3%, nearly one in three lymphoma cases missed by both assessors simultaneously, is striking and clinically important. These figures align with the broader orbital biopsy literature: Koukkoulli *et al.*, reporting from this same institution over the earlier period 2003-2015, found correct pre-biopsy diagnosis in fewer than 50% of all orbital cases, identifying haematological-inflammatory differentiation as the principal challenge [2]. The present study extends those findings by providing formal sensitivity and specificity metrics for the specific lymphoma/inflammation sub-problem, and by characterising anatomical and clinical predictors not previously examined at this centre.

The age ROC AUC of 0.874 represents a strong discriminatory performance. The optimal Youden threshold of age ≥ 55 years, with sensitivity 94.4% and NPV 95.0%, suggests that patient age is arguably the single most powerful pre-biopsy indicator available. However, as this threshold was derived and validated in the same dataset, performance estimates may reflect optimism; external validation is required before clinical adoption of this cut-off. Orbital lymphoma predominantly affects older patients, as reflected in our median age of 74 years for the lymphoma group, consistent with published series reporting median ages in the seventh to eighth decades [9]. Clinicians and radiologists should explicitly factor patient age into their pre-test probability when assessing any orbital mass, and a patient over 55 presenting with a progressive orbital lesion should have lymphoma near the top of their differential regardless of imaging characteristics.

The finding that extraconal lymphoma carries the highest FN-Both rate (44.4%) is a novel contribution of this study. Lacrimal gland masses are well-recognised as a site for orbital lymphoma and are perhaps more readily considered by orbital surgeons; however, our data indicate that lacrimal gland lymphoma carries a substantial missed-diagnosis risk (FN-Both rate 30.0%). The predominant error was misclassification as inflammatory orbital disease. Orbital lymphoma and inflammatory conditions share substantial clinical and radiological overlap, with similar presenting features, comparable mass characteristics on imaging, and no single distinguishing sign, making pre-biopsy differentiation inherently unreliable [3]. Extraconal lesions, by contrast, may be more readily misattributed to inflammatory or myogenic pathology. The extraconal space lacks the anatomical specificity of the lacrimal fossa, and extraconal lymphoma may generate less characteristic MRI signal features, with overlap between lymphoma and idiopathic orbital in-

flammation well documented in the diffusion-weighted imaging literature [10]-[13]. Extraconal location should therefore heighten rather than lower pre-biopsy suspicion for lymphoma.

The significant association of proptosis with lymphoma ($p = 0.015$) is consistent with the known pathophysiology of orbital lymphoma, where the mass effect from relatively rapid-growth lesions produces proptosis more reliably than the more diffuse, perineural, or lacrimal-dominant inflammatory conditions. Pain, although trending towards the inflammatory group, did not achieve significance in this series, likely reflecting limited statistical power rather than a true null effect, and remains clinically useful as a feature that should favour an inflammatory diagnosis.

The counterintuitive finding that lymphoma cases had shorter symptom duration (median 5 months) than inflammatory cases (median 6 months, mean 17.2 months) requires careful interpretation. Orbital lymphoma, particularly EMZL/MALT, which accounted for the majority of lymphoma subtypes in this series, is an indolent malignancy and is not expected to present acutely. The substantially higher mean duration in the inflammatory group reflects the relapsing clinical course of IgG4-related orbital disease, orbital sarcoidosis, and recurrent non-specific orbital inflammation; patients with these conditions may have symptoms for one to two years before a biopsy decision is made.

The four failure patterns identified in FN-Both cases carry direct educational and systems-level implications. Descriptive-only radiology reports represent an avoidable failure mode: a report that provides only dimensional or anatomical data, without offering a differential diagnosis, provides no diagnostic decision support and should be considered a quality gap. Structured radiology reporting with mandatory diagnostic impression fields has been advocated in the orbital biopsy literature and is supported quantitatively by the present data [4] [14]. Prior-diagnosis anchoring represents a well-characterized cognitive bias in diagnostic error science; clinicians should maintain heightened vigilance for patients whose orbital disease is "progressing" from a previously confirmed inflammatory diagnosis, given that transformation to or co-occurrence with lymphoma is recognized. The absent-impression category highlights the particular vulnerability of cases where no clinical and radiological impression is available. This constitutes a quality gap in our practice.

This study has several limitations. Its single-centre retrospective design limits generalisability. The sample size, though the largest dedicated lymphoma/inflammation subgroup series from a UK centre, limits precision in subgroup estimates, particularly for the intraconal compartment ($n = 1$ lymphoma). Pre-biopsy impressions were extracted from existing clinical documentation rather than structured assessments, reflecting real-world practice but introducing potential recording bias. The 2014-2015 overlap with the previously published institutional audit means a small number of cases may appear in both studies, though the focused sub-disease question and novel clinicopathological variables constitute independent original analysis. Prospective studies incorporating advanced imaging modal-

ities, particularly diffusion-weighted MRI, which has demonstrated AUCs exceeding 0.93 in dedicated imaging series, may achieve higher pre-biopsy accuracy than the standard clinical/radiological impression evaluated here [12].

5. Conclusion

Pre-biopsy clinical and radiological assessment at a UK tertiary hospital reliably raises suspicion of orbital lymphoma when positive but cannot safely exclude it when negative. Age over 55 years is a powerful discriminator (AUC 0.874) and should be weighted heavily in pre-test probability estimation. Extraconal anatomical location is specifically associated with the highest risk of missed lymphoma diagnosis. Proptosis favors a diagnosis of lymphoma, while a shorter-than-expected symptom duration does not argue against it. Descriptive-only radiology reports without interpretive diagnostic impression represent a systematic and avoidable failure mode. Orbital biopsy remains the indispensable reference standard for orbital lesions in which lymphoma is a clinical possibility, regardless of pre-biopsy clinical or radiological assessment.

Financial Disclosure

The authors have no relevant financial disclosures for this study.

Ethics Statement

This study was conducted as a retrospective clinical audit at Leeds Teaching Hospitals NHS Trust in accordance with institutional governance requirements. All data were de-identified prior to analysis. Individual patient consent was not required for this audit.

Acknowledgements

The authors thank the histopathology and radiology departments at Leeds Teaching Hospitals NHS Trust for their contribution to the clinical care of the patients included in this series.

Author Contributions

Chow, K.M. collected, analysed the data and drafted the manuscript. Chang, B., Kalantzis, G., Jyothi, S., El-Hindy, N. and Guevara, G. contributed to surgical care. Guevara, G. also contributed to study design and critical revision. Senthilkumar, N. and Arzikulova, B. collected data. All authors approved the final version.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Shields, J.A., Shields, C.L. and Scartozzi, R. (2004) Survey of 1264 Patients with Or-

bital Tumors and Simulating Lesions: the 2002 Montgomery Lecture, Part 1. *Ophthalmology*, **111**, 997-1008. https://doi.org/10.1016/j.ophtha.2003.01.002

[2] Koukkoulli, A., Pilling, J.D., Patatas, K., El-Hindy, N., Chang, B. and Kalantzis, G. (2018) How Accurate Is the Clinical and Radiological Evaluation of Orbital Lesions in Comparison to Surgical Orbital Biopsy? *Eye*, **32**, 1329-1333. https://doi.org/10.1038/s41433-018-0078-3

[3] Ben Simon, G.J., Annunziata, C.C., Fink, J., Villablanca, P., McCann, J.D. and Goldberg, R.A. (2005) Rethinking Orbital Imaging: Establishing Guidelines for Interpreting Orbital Imaging Studies and Evaluating Their Predictive Value in Patients with Orbital Tumors. *Ophthalmology*, **112**, 2196-2207. https://doi.org/10.1016/j.ophtha.2005.09.013

[4] Chen, Y., Hodgkinson, W., Charlton, F., Ting, D.S.J., Meredith, P. and Clarke, L. (2026) A 20-Year Review of Orbital Biopsy at the Newcastle Eye Centre. *Eye*, **40**, 39-44. https://doi.org/10.1038/s41433-025-04079-4

[5] Tang, A., Gout, T., Chang, B., El-Hindy, N. and Kalantzis, G. (2023) Accuracy between Clinical and Radiological Diagnoses Compared to Surgical Orbital Biopsies. *International Journal of Ophthalmology*, **16**, 616-622. https://doi.org/10.18240/ijo.2023.04.16

[6] Kneafsey, S., MacSwiney, T., McCloskey, C., O'Keane, C. and Fulcher, T. (2023) Clinical Characteristics, Diagnosis, and Outcomes of Orbital Biopsies in a Single Irish Centre. *Irish Journal of Medical Science* (1971-), **192**, 1059-1064. https://doi.org/10.1007/s11845-022-03133-4

[7] Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L., *et al.* (2015) STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *BMJ*, **351**, h5527. https://doi.org/10.1136/bmj.h5527

[8] Swerdlow, S.H., Campo, E., Pileri, S.A., Harris, N.L., Stein, H., Siebert, R., *et al.* (2016) The 2016 Revision of the World Health Organization Classification of Lymphoid Neoplasms. *Blood*, **127**, 2375-2390. https://doi.org/10.1182/blood-2016-01-643569

[9] Demirci, H., Shields, C.L., Shields, J.A., Honavar, S.G., Mercado, G.J. and Tovilla, J.C. (2002) Orbital Tumors in the Older Adult Population. *Ophthalmology*, **109**, 243-248. https://doi.org/10.1016/s0161-6420(01)00932-0

[10] Cytryn, A.S., Putterman, A.M., Schneck, G.L., Beckman, E. and Valvassori, G.E. (1997) Predictability of Magnetic Resonance Imaging in Differentiation of Orbital Lymphoma from Orbital Inflammatory Syndrome. *Ophthalmic Plastic & Reconstructive Surgery*, **13**, 129-134. https://doi.org/10.1097/00002341-199706000-00007

[11] Kapur, R., Sepahdari, A.R., Mafee, M.F., Putterman, A.M., Aakalu, V., Wendel, L.J.A., *et al.* (2009) MR Imaging of Orbital Inflammatory Syndrome, Orbital Cellulitis, and Orbital Lymphoid Lesions: The Role of Diffusion-Weighted Imaging. *American Journal of Neuroradiology*, **30**, 64-70. https://doi.org/10.3174/ajnr.a1315

[12] Sun, B., Song, L., Wang, X., Li, J., Xian, J., Wang, F., *et al.* (2016) Lymphoma and Inflammation in the Orbit: Diagnostic Performance with Diffusion-Weighted Imaging and Dynamic Contrast-Enhanced MRI. *Journal of Magnetic Resonance Imaging*, **45**, 1438-1445. https://doi.org/10.1002/jmri.25480

[13] Günay Bulut, E., Gümeler, E., Karakaya, J., Kılıç, L., Karadağ, Ö. and Karlı Oğuz, K. (2026) T2 Signal Ratio Enhances the Diagnostic Performance of Apparent Diffusion Coefficient in Differentiating Orbital Lymphoma from Inflammatory Mimickers. *Diagnostic and Interventional Radiology*, **32**, 457-462. https://doi.org/10.4274/dir.2025.253536

[14] Hunt, S.V., Pereni, I., Williams, M., Ford, R. and Garrott, H. (2022) Multi-Discipli-

nary Team Meetings with Specialist Radiologists May Improve Pre-Operative Clinico-Radiological Diagnostic Accuracy in Patients Requiring Orbital Biopsy and Offer Reciprocal Educational Opportunities. *Eye*, **36**, 2200-2204.
https://doi.org/10.1038/s41433-021-01834-1