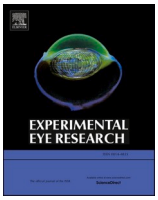
Experimental Eye Research 271 (2026) 111159



Contents lists available at ScienceDirect

Experimental Eye Research

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



Letter to the Editor

Comments on Xu et al. (2026) "Development of a microscopy-based diagnostic test for alkali injury–induced limbal stem cell deficiency through autofluorescence multispectral imaging" (Exp. Eye Res. 264:110829)

Xu et al. present a microscopy-based diagnostic framework for alkali injury-induced limbal stem cell deficiency (LSCD) that combines autofluorescence multispectral imaging (AFMI) with statistical preprocessing and supervised machine-learning classification (Xu et al., 2026; Exp. Eye Res. 264:110829). By integrating multispectral feature extraction with a classification pipeline, they demonstrate high diagnostic accuracy across two imaging platforms, achieving 90.38% accuracy on the Operetta CLS system and 83.28% on the Olympus IX83, with k-nearest neighbors (KNN) serving as a nonparametric classifier for separating LSCD-affected from healthy corneal epithelium. The reported concordance between AFMI-based predictions and histopathological assessment indicates that this approach may offer an objective strategy for LSCD diagnosis. However, the analytical workflow—centered on univariate statistical filtering operating under implicit independence and distributional assumptions, together with linear dimensionality reduction—warrants further discussion regarding its ability to capture potentially nonlinear structure in multispectral biological data.

Initial feature selection relies on univariate statistical testing to identify spectral channels with significant differences between healthy and LSCD tissues. While such screening is appropriate for dimensionality reduction and noise exclusion, it operates under simplifying assumptions, including approximate independence and distributional forms, and treats each channel independently. Autofluorescence signals from NADH, FAD, and PpIX arise from biologically coupled metabolic and redox processes that respond coordinately to epithelial injury and remodeling, such that treating these channels as independent observables may obscure higher-order dependencies relevant to LSCD progression and heterogeneity—an issue recognized in analyses relying on linear statistical assumptions.

They employ Linear Discriminant Analysis (LDA) as a supervised dimensionality-reduction step prior to classification. LDA is a linear, parametric approach widely used in biological assay analysis for interpretability and computational efficiency under standard distributional assumptions (Jarantow et al., 2023; Curr. Protoc. 3:e801). They note that reducing the feature space to 10 dimensions for the Operetta CLS system or 20 dimensions for the Olympus IX83 platform represents a practical compromise among accuracy, generalizability, and efficiency. At the same time, it should be noted that the use of LDA as an intermediate representation imposes a linear structure on the feature space, prioritizing variation preserved under this projection and potentially reducing sensitivity to complex feature interactions that may reflect heterogeneous LSCD states.

Final classification is carried out using KNN, a supervised, nonparametric method capable of modeling non-linear decision boundaries. While KNN itself does not impose distributional assumptions, its predictive behavior depends on the geometry of the feature space. When applied downstream of linear dimensionality reduction, classification performance may reflect properties of the imposed linear embedding rather than intrinsic non-linear relationships in the original data. Although linear projections are widely used in biomedical imaging and can yield meaningful separation (Caprihan et al., 2008; NeuroImage 42:675–682), they may also introduce biases through projection-dependent variance structure (Dey and Lee, 2019; J. Multivar. Anal. 173:145–164). Although overfitting is addressed through cross-validation and mouse-level masking, sector-level classification may still be influenced by local spatial correlations within individual corneas.

Principal Component Analysis (PCA) is also used exclusively for visualization and exploratory interpretation. Although PCA summarizes global variance patterns, PCA-based interpretation is sensitive to linear variance structure and may be biased when used to infer biological relationships (Elhaik, 2022; Sci. Rep. 12:14683), with additional limitations noted in genetic and population-level contexts (Yao and Ochoa, 2023, eLife 12:e79238). Recent analyses further suggest that PCA-derived patterns may, in some cases, reflect projection-dependent structure rather than fully representing underlying biological proximity, warranting cautious interpretation of exploratory visualizations. These variance-based projections, while effective for reducing dimensionality and suppressing noise, rely on simplifying assumptions, including approximate independence and specific distributional forms across spectral channels. In this context, it should be noted that PCA in this study is used solely for exploratory visualization and is not involved in the classification step, and is not intended to provide definitive inference regarding underlying biological relationships.

Taken together, combining univariate statistical filtering operating under implicit independence and distributional assumptions with linear dimensionality reduction provides an interpretable analytical framework but may limit sensitivity to non-linear structure in multispectral autofluorescence data. Interpretability itself, however, is a nuanced and context-dependent concept in machine learning, and should therefore be interpreted with appropriate consideration of the underlying modeling assumptions (Lipton, 2018; ACM Queue 16:31–57). This limitation does not undermine the reported classification performance supported by blinded testing and agreement with histology, but contextualizes conclusions regarding the organization and separability of LSCD states as conditioned on linear representations of correlated features, noting that predictive performance does not necessarily guarantee robustness of feature-level interpretations (Fisher et al., 2019; JMLR 20:177). To address these limitations within the reported framework, complementary analyses could be considered. Feature Agglomeration may enable

https://doi.org/10.1016/j.exer.2026.111159
Received 1 May 2026; Received in revised form 2 July 2026; Accepted 6 July 2026
Available online 10 July 2026
0014-4835/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

exploration of higher-level structure among correlated spectral channels without enforcing linearity or orthogonality constraints (Zhang et al., 2020; Neurocomputing 380:180–189). Feature-centric strategies inspired by Highly Variable Gene Selection could further reduce projection dependence while emphasizing coordinated responses associated with LSCD severity (Xie et al., 2025; BMC Bioinformatics 26:104). In addition, rank-based nonparametric association measures (Spearman's rho, Kendall's tau) could quantify condition-dependent feature couplings without strong distributional assumptions (Okoye and Hosseini, 2024, in: R Programming: Statistical Data Analysis in Research, Springer; Yu and Hutson, 2024, Commun. Stat. Theory Methods 53:2141–2153). It is important to recognize that the sample size in the original study may limit the direct application of more complex nonlinear approaches, and the suggestions presented here are intended to be complementary rather than replacements. This perspective is further supported by the linear and parametric nature of PCA-based representations, which may benefit from complementary analytical perspectives when exploring complex feature relationships.

In conclusion, Xu et al. provide a well-validated AFMI-based diagnostic framework for LSCD that integrates imaging, statistics, and machine learning. Explicit recognition of the linear and parametric constraints embedded in feature selection, dimensionality reduction, and visualization highlights opportunities for complementary nonparametric analyses aimed at better capturing the complex biology of corneal epithelial injury and repair. Incorporating such perspectives may further enhance robustness and generalizability as AFMI-based diagnostics advance toward broader experimental and clinical application.

CRediT authorship contribution statement

Souichi Oka: Conceptualization, Writing – original draft. **Ryota Ono:** Investigation. **Yoshiyasu Takefuji:** Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Souichi Oka[a,*], Ryota Ono[b], Yoshiyasu Takefuji[b]
[a] *Faculty of Data Science, Musashino University, 3-3-3 Ariake, Koto-ku, Tokyo, 135-8181, Japan*
[b] *Science Park Corporation, 3-24-9 Iriya-Nishi, Zama-shi, Kanagawa, 252-0029, Japan*

[*] Corresponding author.
E-mail addresses: s-oka@musashino-u.ac.jp (S. Oka), rono@sciencepark.co.jp (R. Ono), takefuji@keio.jp (Y. Takefuji).