Received: 3 October 2023 | Accepted: 3 March 2024

DOI: 10.1111/opo.13302

ORIGINAL ARTICLE

OPO THE COLLEGE OF OPTOMETRISTS

Machine learning-based prediction of tear osmolarity for contact lens practice

Izabela K. Garaszczuk[1] | **Maria Romanos-Ibanez**[2] | **Alejandra Consejo**[2]

[1]Wroclaw University of Science and Technology, Wroclaw, Poland

[2]Aragon Institute for Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain

Correspondence
Alejandra Consejo, Department of Applied Physics, University of Zaragoza, Zaragoza, Spain.
Email: alejandra.consejo@unizar.es

Funding information
Cátedra SAMCA de Desarrollo Tecnológico de Aragón, Universidad de Zaragoza (III Premio a la Innovación Multidisciplinar); MCIN/AEI/10.13039/501100011033 and European Union "NextGenerationEU"/PRTR, Grant/Award Number: TED2021-130723A-I00

Abstract

Purpose: This study addressed the utilisation of machine learning techniques to estimate tear osmolarity, a clinically significant yet challenging parameter to measure accurately. Elevated tear osmolarity has been observed in contact lens wearers and is associated with contact lens-induced dry eye, a common cause of discomfort leading to discontinuation of lens wear.

Methods: The study explored machine learning, regression and classification techniques to predict tear osmolarity using routine clinical parameters. The data set consisted of 175 participants, primarily healthy subjects eligible for soft contact lens wear. Various clinical assessments were performed, including symptom assessment with the Ocular Surface Disease Index and 5-Item Dry Eye Questionnaire (DEQ-5), tear meniscus height (TMH), tear osmolarity, non-invasive keratometric tear film break-up time (NIKBUT), ocular redness, corneal and conjunctival fluorescein staining and Meibomian glands loss.

Results: The results revealed that simple linear regression was insufficient for accurate osmolarity prediction. Instead, more advanced regression models achieved a moderate level of predictive power, explaining approximately 32% of the osmolarity variability. Notably, key predictors for osmolarity included NIKBUT, TMH, ocular redness, Meibomian gland coverage and the DEQ-5 questionnaire. In classification tasks, distinguishing between low (<299 mOsmol/L), medium (300–307 mOsmol/L) and high osmolarity (>308 mOsmol/L) levels yielded an accuracy of approximately 80%. Key parameters for classification were similar to those in regression models, emphasising the importance of NIKBUT, TMH, ocular redness, Meibomian glands coverage and the DEQ-5 questionnaire.

Conclusions: This study highlights the potential benefits of integrating machine learning into contact lens research and practice. It suggests the clinical utility of assessing Meibomian glands and NIKBUT in contact lens fitting and follow-up visits. Machine learning models can optimise contact lens prescriptions and aid in early detection of conditions like dry eye, ultimately enhancing ocular health and the contact lens wearing experience.

KEYWORDS
contact lenses, dry eye disease, machine learning, regression and classification techniques, tear osmolarity

INTRODUCTION

Machine learning methods have seen widespread adoption in various domains, including ophthalmology,[1] and yet their application in soft contact lens research remains underutilised. This study addresses the application of machine learning techniques in estimating a clinically important, however costly, and challenging-to-measure parameter, namely tear osmolarity.

Tear osmolarity, which quantifies the concentration of osmotically active particles in tears, is typically defined as the number of osmoles per litre of solution (mOsm/L). In tears, it is predominantly influenced by the electrolytes found in their aqueous component, with proteins and sugars playing a comparatively lesser role.[2] Tear osmolarity is a valuable clinical metric that provides insight into tear production, evaporation, drainage and absorption.[3] Several studies have suggested that tear osmolarity serves as the most effective solitary diagnostic test for dry eye disease (DED), and serves as an objective numerical gauge for diagnosing, grading the severity of and managing this condition.[4,5] Therefore, it is included in the definition of the disease, and is a part of the recommended diagnostic protocol.[6,7]

As many as 150 million contact lens users worldwide report symptoms of ocular discomfort to be the main reason for contact lens discontinuation.[8,9] While some subjects suffer from contact lens discomfort that subsides after lens removal, contact lens-induced DED does not always resolve after cessation of contact lens use. It has been suggested that elevated tear osmolarity is a key feature in contact lens-induced DED.[10,11] This makes tear hyperosmolarity a critical factor in the development of DED in individuals—both contact lens wearers and non-wearers.[12]

Consequently, effectively managing tear osmolarity becomes essential for ensuring the proper fit of contact lenses. Tear hyperosmolarity has been observed in individuals who wear both soft and rigid contact lenses, whether on a daily or extended-wear basis.[13] Research indicates that the increase in tear osmolarity during contact lens wear can be attributed to several factors, including environmental conditions, tear film properties, contact lens materials and specifications, as well as the wearing schedule.[14–16] In a previous study, it was demonstrated that contact lens users with higher initial tear osmolarity rates may benefit from switching to daily disposable contact lenses, resulting in reduced contact lens discontinuation rates.[17] The identification of such individuals in need of special attention is of clinical importance.

Despite all the above-mentioned benefits of including tear osmolarity in the clinical protocol, its utility is hampered by the cost and complexity associated with even the simplest point of care devices such as the Tear Lab Osmolarity system (tearlab.com). While the accuracy of tear osmolarity measurements performed with this device remains consistent across all osmolarity ranges, in vivo measurements displayed poor repeatability, which puts doubt on the diagnosis when only one sample is taken from each eye.[18] This requires the use of more than one lab-on-chip for sampling, making it more costly and time consuming. As a result, the widespread assessment of osmolarity remains limited in routine clinical practice, especially in the assessment of contact lens fit, and is mostly confined to the research setting.[13]

This study aimed to explore various techniques for predicting osmolarity using some routine clinical parameters that could be assessed using standard clinical tools for contact lens fit and ocular surface assessment.

Key points

- Machine learning models can provide a non-invasive and inexpensive tool to optimise contact lens prescriptions and aid in early detection of conditions like dry eye disease.
- This study demonstrated the limitations of simple linear regression for osmolarity prediction and underscores the need for advanced machine learning models in clinical settings.
- This study highlights the importance of assessing tear film stability and Meibomian gland morphology for precise tear osmolarity prediction and improved contact lens fitting.

METHODS

Subjects and data collection

This study was performed on a population of healthy subjects who were considered likely to be successful contact lens wearers. The study pool comprised individuals at least 18 years of age who were either habitual contact lens wearers fitted with contact lenses by the (IKG) of this study or who had expressed an interest in wearing contact lenses. Participants had the opportunity to enrol through an online registration form. A total of 175 participants (175 eyes) were included in this study. Some of the data were retrospectively collected from *baseline* measurements from a previous study of tear osmolarity in contact lens wearers.[17] The study adhered to the tenets of the Declaration of Helsinki and was approved by the University's ethical board. Written informed consent was obtained from all participants. Participants were adults (133 females, 42 males) between 20 and 77 years old (mean: 31 ± 13 years old). Exclusion criteria included the absence of severe dry eye, significant inflammation or substantial impairment of tear flow, as well as the absence of any systemic conditions that were recognised for their potential to negatively affect the ocular surface and therefore impact soft contact lens wear. Some subjects were contact lens wearers

before commencing the study and were asked to refrain from wearing contact lenses at least 3 days prior to their participation. As shown in a previous investigation, a 3-day break should be sufficient to stabilise the tear film before commencing the study.[17]

Study protocol

The study protocol included qualifying subjects for contact lens wear. Measurements were performed from the least to the most invasive.[19] Firstly, after a medical history check, the Ocular Surface Disease Index (OSDI) and 5-Item Dry Eye Questionnaire (DEQ-5) were administered to assess ocular surface and DED symptoms. Both questionnaires were administered in the subjects' native language.[20] While the OSDI is more focused on symptoms of any ocular surface disease, the DEQ-5 was used to rule-out symptoms specific for DED. Measurements described below were performed by the same practitioner under consistent conditions (in the same laboratory with controlled temperature and relative humidity). For osmolarity measurements, the same diagnostic pen was used each time and care was taken to sample tears from the same area of the tear meniscus.

Ocular symptom assessment was followed by measurement of the tear meniscus height (TMH) with the Oculus Keratograph 5M (K5M; Oculus Optikgeräte GmbH, oculus.de).[21] Secondly, tear osmolarity was measured using an in vitro diagnostic device (TearLab™ Osmolarity System, tearlab.com). This osmometer is a disposable 'lab-on-a-chip' system, which requires less than 50 nanolitres of tear fluid to be sampled for analysis.[22] A desktop instrument transforms the electrical signals produced by the laboratory card into a numerical measurement and presents the result on a screen. After calibration of the instrument at the beginning of each session, tears were collected by the same trained investigator with a single-use test card directly from the lower temporal part of the tear meniscus. The device immediately displays the osmolarity value in mOsm/L after docking the probe with a chip in the device. The values were documented. This process involved two to three measurements for each eye, and the average value was recorded. If the two measurements were not consistent (differed more than 10 mOsm/L), a third measurement was performed and the most outlying value was omitted.

Subsequently, the non-invasive Keratograph break-up time (NIKBUT) was assessed with the K5M.[23,24] This device measures tear film break-up based on the quality of the images of Placido disks reflected from the surface of the tear film. As described in a previous study,[25] two readings were taken at the end of each NIKBUT assessment; the First NIKBUT, which was the time taken from a blink to the first appearance of substantial deformation of the reflected image that indicates tear film break-up, and the Mean NIKBUT, which was the average time taken from the blink to the ring deformations in all regions monitored over the duration of the recording. Each recording lasted a maximum of 24 s.

Afterwards, ocular limbal and bulbar redness was scored using a K5M function called R-Scan. The temporal and nasal values of limbal redness and the temporal and nasal values of bulbar redness were averaged. The K5M used bright visible white light to score ocular redness, which results in some subjects reporting light sensitivity and tearing. Therefore, this procedure was performed after NIKBUT estimation, so as not to influence the tear film.

Further, slit-lamp fluorescein staining scoring with the Efron scale was performed (Righton RS1000, righton-oph.com) with a cobalt blue filter and Wratten 12 yellow filter after one drop of 0.9% saline solution was used to moisten a 1 mg fluorescein sodium ophthalmic sterile strip (BioGlo; HUB Pharmaceuticals Inc, hubrx.com) and administered to the lower conjunctival sac.[26]

Lastly, the Meibomian glands were visualised in infrared light with the Meibo-Scan K5M tool.[27,28] Default settings were used with 0.5× magnification. Images of both the lower and upper everted eyelids were acquired and saved for further processing. ImageJ (US National Institutes of Health, imagej.nih.gov) was used to perform the image processing. The Polygon selection tool was used to mark and calculate the surface area of two regions: the area of the eyelid where Meibomian glands were not present (drop-out area) and the area of the whole exposed eyelid. The Meibomian Upper and Lower Lid Percentages were calculated as the percentage of the drop out area compared with the whole exposed surface of the everted eyelid (upper and lower eyelid, respectively).

Data pre-processing and machine learning regression and classification algorithms were implemented in Python (Python Software Foundation, python.org) and were based on the scikit-learn data science library (scikit-learn.org).

Variable selection for osmolarity prediction

Based on the protocol described above, a set of 11 input parameters was used to predict osmolarity. These independent variables, serving as predictors for the dependent variable (osmolarity), included first NIKBUT, mean NIKBUT, TMH, bulbar ocular redness, limbal ocular redness, fluorescein corneal staining, fluorescein conjunctival staining, Meibomian upper lid percentage, Meibomian lower lid percentage, OSDI and DEQ-5 score. To gain insights into potential interrelationships between these variables, their pairwise linear correlations were evaluated using the Pearson coefficient (r), which ranges from −1 to +1. These correlations were visualised in a correlation heat map, a graphical tool that presents the relationships between multiple variables as a colour-coded matrix.

Variable selection in the training of regression and classification models is a crucial step in determining which independent variables are included in the model. While some algorithms automatically choose the best combination of

variables, others provide the operator with more flexibility. In such cases, the selection process can follow two primary approaches: forward selection, where variables are added one by one, starting with none and including the most relevant at each step, and backward selection, where all variables initially serve as candidates, and the least important ones are iteratively eliminated. Both methods were employed when applicable.

Data pre-processing

Data pre-processing is essential before testing any regression or classification machine learning model because it helps ensure that the input data is clean, consistent and appropriately formatted, improving the model's performance, reducing the risk of overfitting and enhancing the interpretability of the results.[1] Data pre-processing consisted of four steps described below.

1. Imputation: This statistical technique is used in data analysis and pre-processing to fill in missing or incomplete data with estimated values. Imputation is an essential step in data pre-processing, as missing data can lead to biased or inaccurate results in machine learning models if not handled properly. *K*-nearest neighbours (KNN), with $k=5$, was chosen as the imputation technique in variables with missing values. KNN is often better than linear correlation for imputation because it considers a broader range of similar data points, making it more robust for handling complex relationships and non-linear patterns in the data.
2. Outlier detection: Outlier detection helps identify data points that deviate significantly from the norm, which can profoundly impact model accuracy and the validity of the insights drawn from the data. The interquartile range (IQR) outlier detection method was applied. It involves calculating the range between the first quartile (25th percentile) and the third quartile (75th percentile) of a data set. Data points that fall below the first quartile minus 1.5 times the IQR or above the third quartile plus 1.5 times the IQR are considered as potential outliers and often flagged for further investigation or treatment in data analysis. The IQR method was chosen for its simplicity and robustness. IQR is robust to the presence of extreme outliers because it is based on the quartiles of the data rather than the mean and standard deviation, which can be heavily influenced by outliers. From the 175 available instances, 164 remained after outlier depletion (i.e., 6.3% of the whole data set, which was considered acceptable).
3. Data scaling: Data scaling is important in pre-processing because it ensures that variables with different scales do not unduly influence machine learning algorithms, enabling fair comparisons between features and preventing certain algorithms from being dominated by the magnitude of the data. It also helps improve convergence and performance in several machine-learning models. *Z*-score standardisation, also known as standard score transformation, was chosen. This is a data scaling technique that involves transforming data into a standardised scale with a mean of zero and a standard deviation of one by subtracting the mean and dividing by the standard deviation of the data. Advantages of *Z*-score standardisation over other data scaling methods include making data interpretable in terms of standard deviations from the mean, allowing for easy comparison of variables with different units and scales and reducing the sensitivity of models to outliers as extreme values are brought within a similar range.
4. Dividing the data set in training and validation: The test set (validation) is a distinct portion of the data set that remains unused during the training phase. Instead, it comes into play after the model has been trained to assess its performance and its ability to generalise new data. The test set offers a realistic evaluation of the model's accuracy and effectiveness in real-world scenarios. It is crucial for the test set to be as independent as possible from the training set to avoid overfitting. The process of data splitting was executed randomly to prevent biases in both the training and test sets. For regression models, allocating 80% of the data for training and the remaining 20% for validation is common. In the case of classification models, a split of 75% for training and 25% for validation is often chosen. These were the chosen proportions in this work. The variation in data proportion is based on the specific characteristics of each problem type. In regression, the output is a continuous numerical value, necessitating the model to learn more intricate relationships for precise predictions. Conversely, in classification, the task involves assigning discrete labels or categories, allowing models to learn more swiftly, as the relationships between classes are typically more distinct and less subtle in most cases. Consequently, in classification problems, allocating a relatively smaller portion of data for training is feasible while still achieving satisfactory results. However, in regression problems, having a larger volume of training data is essential for the model to capture adequately the intricacies of the continuous values it aims to predict.

Machine learning regression algorithms

Four supervised regression algorithms were implemented and tested: (1) multiple linear regression, (2) polynomial regression, (3) support vector regression (SVR) and (4) random forest regression. Within a given regression algorithm, certain setting variations were allowed. In this work, different settings were tested, resulting in the selection of algorithms due to their higher performance. Specifically, a second-order polynomial regression and SVR using a polynomial kernel with a power of three was

implemented. The coefficient of determination (R^2) was used to assess the performance of different regression models. R^2 ranges from 0 to 1 and represents the proportion of variance in the dependent variable (osmolarity) explained by the independent variables in a regression model, where 0 indicates no explanatory power and 1 indicates a perfect fit.

Machine learning classification algorithms

Six supervised classification algorithms were implemented and tested: (1) Logistic regression, (2) KNN, (3) Support vector machine (SVM), (4) Naïve Bayes, (5) Decision tree and (6) Random forests. Similar to regression algorithms, some classification algorithms (such as KNN and SVM) offer certain degrees of flexibility. The best performance was achieved using KNN with seven neighbours and SVM with a radial kernel (rbf). The confusion matrix for the training data set was calculated to evaluate the performance of different classification models. The confusion matrix provides information on true positives, true negatives, false positives and false negatives. With these values, accuracy, precision, sensitivity and specificity were computed.

The DED status in an individual in terms of tear osmolarity was measured against values derived from population norms, generated from subjects of both sexes representing a wide age range. Average tear osmolarity in healthy adults is around 302 ± 9.7 mOsm/L.[7,29,30] Jacobi et al. using the TearLab osmometer reported values around 301 mOsm/L, ranging from 298 to 304 mOsm/L.[31,32] Since this study included mainly heathy subjects who qualified for soft contact lens wear classifications, the focus was on normal osmolarity values. Therefore, in classification algorithms, two cases were investigated. Case 1 consisted of a binary classification, that is, two groups: osmolarity <300 mOsmol/L ($n = 94$) and osmolarity >300 mOsmol/L ($n = 70$). Case 2 consisted of a tripartite classification, that is, three groups, defined as low osmolarity (<299 mOsmol/L [$n = 94$]), medium osmolarity (300–307 mOsmol/L [$n = 80$]) and high osmolarity (>308 mOsmol/L [$n = 90$]). The osmolarity of 308 mOsm/L was considered to be the most sensitive threshold to distinguish normal from mild/moderate forms of DED.[5] In tripartite classification, random noise was applied as a data augmentation technique in the medium and high osmolarity groups. Data augmentation techniques for numerical variables revolve around introducing controlled randomness or variability into the data to create additional training samples. Given the constraints of the data set, which comprised 164 instances, expanding beyond this tripartite classification was carefully considered but ultimately not pursued. Introducing more classes in a data set of this size could significantly increase the risks of bias and noise, specifically leading to overfitting, class imbalance and noise amplification.

RESULTS

Pairwise linear correlations

As depicted in Figure 1, none of the 11 predictor variables exhibited a direct relationship with osmolarity (all $p > 0.05$). The strongest correlations were found between similar parameters, that is, first and mean NIKBUT ($r = 0.91$, $p < 0.001$), bulbar and limbal ocular redness ($r = 0.32$, $p < 0.001$), upper and lower Meibomian lid coverage ($r = 0.41$, $p < 0.001$) and OSDI and DEQ-5 questionnaires ($r = 0.60$, $p < 0.001$).

Machine learning regression algorithms

The most successful regression model, as shown in Table 1, achieved an R^2 of 0.32 using multiple linear regression. This suggests a moderate level of predictive power, indicating that approximately 32% of the variability in osmolarity can be explained by the included variables in the model, namely NIKBUT, TMH, ocular limbal redness, Meibomian upper lid coverage and the DEQ-5 questionnaire.

Machine learning classification algorithms

In binary classification (Case 1), the highest accuracy achieved in distinguishing between osmolarity groups (<300 and >300 mOsmol/L) was 78% using decision trees, as presented in Table 2. Notably, this model employed the same input variables as the multiple regression model presented in Table 1; specifically: NIKBUT, TMH, ocular limbal redness, Meibomian upper lid coverage and the DEQ-5 questionnaire.

In tripartite classification (Case 2), the highest accuracy attained for distinguishing between low, medium or high osmolarity levels was 83%, employing random forests, as detailed in Table 3. The input variables for this model closely resembled those used in the binary classification algorithm, with the primary distinction being the absence of TMH. Remarkably, the input variables for both binary and tripartite classification remained consistent across each respective model.

DISCUSSION

This study explored various techniques for predicting osmolarity using routine clinical parameters. The findings indicate that simple linear regression is inadequate for accurate osmolarity prediction, as evident from Figure 1. Therefore, the task of predicting osmolarity necessitates more advanced regression models. The regression models employed in the current analysis achieved a performance of 32% (Table 1), while classification models

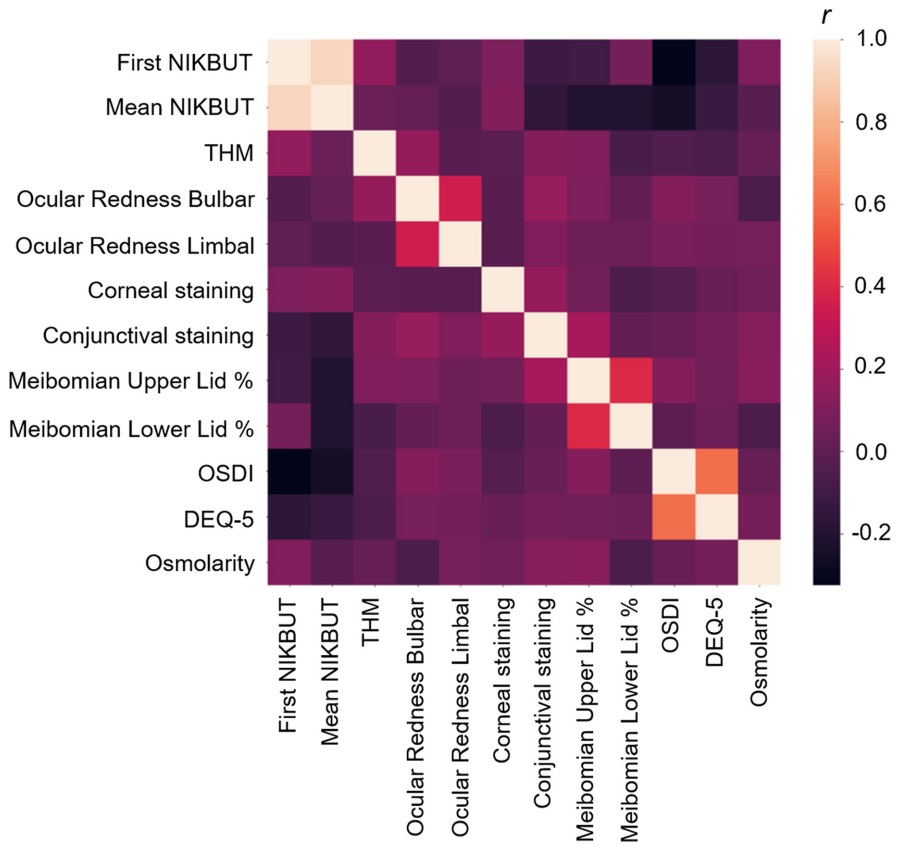

FIGURE 1 Correlation heat map pairwise linear correlations among the study variables evaluated using the Pearson coefficient (*r*). DEQ-5, 5-item dry eye questionnaire; NIKBUT, non-invasive keratometric tear film break-up time; OSDI, ocular surface disease index; THM, tear meniscus height.

TABLE 1 Top-performing regression models for osmolarity estimation, with the model's input variables denoted by (×).

	Multiple lineal regression	Polynomial regression (second degree)	SVR (third-degree polynomial kernel)	Random forests
First NIKBUT	×	×	×	×
Mean NIKNUT	×	×	×	
TMH	×	×		
Ocular bulbar redness				
Ocular limbal redness	×			×
Corneal staining		×		
Conjunctival staining		×		
Meibomian upper lid %	×			
Meibomian lower lid %			×	×
OSDI				
DEQ-5	×		×	
R^2	0.32	0.28	0.23	0.24

Abbreviations: DEQ-5, dry eye questionnaire; NIKBUT, non-invasive keratometric tear film break-up time; OSDI, ocular surface disease index; SVR, support vector regression; TMH, tear meniscus height.

achieved an accuracy of approximately 80% (Tables 2 and 3).

It is important to note that regression and classification serve distinct objectives, contributing to differences in their complexity and performance. Regression seeks to predict the precise numerical value of osmolarity for an individual, while classification focuses on categorising individuals into groups, such as low or high osmolarity, without providing specific osmolarity values. Remarkably, despite variations in model objectives and

TABLE 2 Performance of classification algorithms in binary classification, specifically in distinguishing between osmolarity groups (<300 and >300 mOsmol/L), with the model's input variables marked by (×).

	Logistic regression	KNN ($k=7$)	SVM (rbf kernel)	Naïve Bayes	Decision trees	Random forests
First NIKBUT	×	×		×	×	×
Mean NIKNUT			×	×	×	×
TMH	×				×	
Ocular bulbar redness		×				×
Ocular limbal redness			×		×	
Corneal staining						
Conjunctival staining	×		×			
Meibomian upper lid %	×	×	×		×	×
Meibomian lower lid %		×	×			×
OSDI						
DEQ-5	×		×	×	×	×
Sensitivity	58%	68%	53%	53%	89%	53%
Specificity	82%	82%	91%	82%	68%	95%
Accuracy	71%	76%	73%	68%	78%	76%
Precision	73%	76%	83%	71%	71%	91%

Abbreviations: DEQ-5, dry eye questionnaire; KNN, *K*-nearest neighbours; NIKBUT, non-invasive keratometric tear film break-up time; OSDI, ocular surface disease index; SVM, support vector machine; TMH, tear meniscus height.

performance, the predictor variables consistently featured in all models (Tables 1–3). Notably, NIKBUT, TMH, limbal ocular redness, Meibomian upper lid percentage and the DEQ-5 emerged as key parameters for osmolarity prediction. These key objective parameters were shown to be increased in severe DED.[28] The adopted approach integrated both objective parameters and subjective assessments, such as the OSDI and DEQ-5 questionnaires, to encompass the full spectrum of clinically available tools. Intriguingly, the DEQ-5 questionnaire proved to be a valuable parameter for osmolarity prediction, whereas the OSDI did not exhibit the same level of predictive utility (Tables 1–3).

There is a case for using the Contact Lens Dry Eye Questionnaire-8 (CLDEQ-8) instead of the DEQ-5, as the former better describes contact lens performance and contact lens-related dry eye, whose underlying causes and symptoms differ from the typical manifestations of DED. However, we need to consider that many of the subjects participating in this study ($n=100$) were novice contact lens users with no prior experience of contact lens use. Therefore, they could not (even retrospectively) complete this questionnaire. The CLDEQ-8 could be administered during a follow-up visit, which while being very important in contact lens fitting, was not the part of this study. Additionally, the universal protocol applied here may be performed as a part of general clinical practice, not only for contact lens fitting.

In this study, the primary aim was to demonstrate the potential benefits of integrating machine learning methodologies into contact lens research. Consequently, the participant pool was limited to healthy subjects, as individuals with severe dry eye symptoms often face challenges in wearing regular soft contact lenses. This inherent limitation added complexity to the prediction task.

In scenarios involving individuals with ocular health issues, it was anticipated that alterations in the overall parameter profiles would facilitate both regression and classification predictions. Future iterations of this research would benefit from including dry eye patients, potentially leading to improved predictive model performance. Given that many ocular surface parameters measured in this study were pertinent to multifocal contact lens fitting and exhibit correlations with age,[27] care was taken to include subjects across a wide age range.

In addressing the choice of analytical tools used in this study, it is important to note why regression and classification models were preferred over deep neural network-based solutions, which are increasingly popular in medical data classification.[1] Deep learning techniques are particularly adept at extracting features directly from image data, a capability that was not required for this research which focused on quantified clinical parameters. A key objective of this study was not only to predict tear osmolarity but also to understand the relative significance of various clinical parameters in these predictions. Deep learning models, despite their powerful predictive abilities, often lack transparency in their decision-making processes, functioning as 'black boxes'. This characteristic of deep learning models poses a challenge in medical research where interpretability is crucial. This approach enabled us to maintain a clear insight into the influence of individual clinical parameters on osmolarity levels, an aspect deemed essential for the objectives of this study. However, it would be interesting

TABLE 3 Performance of classification algorithms in tripartite classification, specifically in distinguishing between low (<299 mOsmol/L), medium (300–307 mOsmol/L) and high osmolarity (>308 mOsmol/L), with the model's input variables marked by (×).

	Logistic regression	KNN ($k=7$)	SVM (rbf kernel)	Naïve Bayes	Decision trees	Random forests
First NIKBUT	×	×		×	×	×
Mean NIKNUT			×	×	×	×
TMH	×				×	
Ocular bulbar redness		×				×
Ocular limbal redness			×		×	
Corneal staining						
Conjunctival staining	×		×			
Meibomian upper lid %	×	×	×		×	×
Meibomian lower lid %		×	×			×
OSDI						
DEQ-5	×		×	×	×	×
Sensitivity						
Low osm.	60%	60%	65%	75%	40%	80%
Medium osm.	21%	47%	32%	26%	79%	79%
High osm.	30%	56%	59%	7%	89%	89%
Specificity						
Low osm.	38%	60%	58%	21%	91%	89%
Medium osm.	57%	73%	78%	52%	78%	98%
High osm.	70%	84%	73%	95%	79%	86%
Accuracy						
Low osm.	38%	43%	45%	36%	67%	76%
Medium osm.	21%	47%	43%	24%	63%	94%
High osm.	53%	79%	70%	67%	80%	83%
Precision						
Overall	36%	55%	53%	33%	71%	83%

Abbreviations: DEQ-5, dry eye questionnaire; KNN, *K*-nearest neighbours; NIKBUT, non-invasive keratometric tear film break-up time; OSDI, ocular surface disease index; osm., osmolarity; SVM, support vector machine; TMH, tear meniscus height.

for future work to explore the potential of deep learning, particularly in the analysis of image data, which could uncover new dimensions in the diagnosis and understanding of tear film-related conditions.[33]

The results obtained here support the clinical assessment of Meibomian glands with non-contact infrared Meibography in contact lens fitting. The principal source of dry eye and discomfort in contact lens wearers is not well identified. However, research suggests that the problem is multifactorial, with a strong link towards physiological changes occurring in the eyelids and particularly the Meibomian glands. Alterations to Meibomian gland morphology and function accompany contact lens wear, especially during the first 2 years of use.[34,35] Even though more prolonged exposure beyond this point does not appear to be associated with further shortening of Meibomian glands,[36] one must consider that these changes most probably do not resolve over time after the cessation of contact lens wear.[35] Moreover, Meibomian gland disfunction seems to be the most prominent cause of evaporative DED.[37]

One could argue that non-contact infrared meibography (NIM), as used in this study, also requires a sophisticated and costly device and is not part of most routine examinations. However, many studies show advantages of NIM over standard lid margin assessment; this method has proven useful in diagnosing non-obvious cases of Meibomian gland dysfunction,[38] could be used to diagnose subtle changes in Meibomian gland morphology in subjects with DED and contact lens wearers[34,39,40] and enables a more objective approach[36,41] compared with the routinely used methods of lid margin assessment. Additionally, devices with a NIM option are multitools, enabling extensive clinical evaluation of those parameters that appeared to be key in osmolarity predictions. Alternatively, they may come as hardware attachments to slit-lamp biomicroscopes that are several times less expensive than osmometers. These devices are used worldwide by practitioners and do not require single-use cartridges and calibration solutions that add to the total cost.

This study also suggests including NIKBUT in contact lens assessment. In a previous study, we showed that it

may be clinically beneficial to include NIKBUT assessment in contact lens fitting and in the follow-up visits to increase the rate of successful fits.[25] NIKBUT can be now assessed non-invasively without the use of fluorescein with many devices used in standard clinical settings. These devices can be used as multitools for both tear film assessment and corneal topography. Traditional objective tests to quantify tear film break-up time with the use of fluorescein are limited by their invasiveness, low repeatability and reproducibility; therefore, non-invasive, objective and automatic tests are recommended.

This research presents valuable clinical utility in the context of contact lens applications. By employing machine learning methods to estimate osmolarity accurately, it advances contact lens research and practice. Clinicians can leverage these models to optimise contact lens prescriptions, ensuring improved comfort and visual outcomes for wearers. Additionally, the research provides potential methods to detect early signs of dry eye among contact lens users, enabling proactive management and protocol design, enhancing overall ocular health and the contact lens wearing experience.

AUTHOR CONTRIBUTIONS

Izabela K. Garaszczuk: Conceptualization (lead); data curation (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); resources (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Maria Romanos-Ibanez:** Data curation (equal); formal analysis (lead); methodology (equal); writing – review and editing (equal). **Alejandra Consejo:** Conceptualization (lead); data curation (equal); formal analysis (equal); funding acquisition (lead); investigation (equal); methodology (equal); project administration (equal); resources (equal); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal).

FUNDING INFORMATION

MCIN/AEI/10.13039/501100011033 and European Union 'NextGenerationEU'/PRTR. Grant Number: TED2021-130723A-I00, and Cátedra SAMCA de Desarrollo Tecnológico de Aragón, Universidad de Zaragoza (III Premio a la Innovación Multidisciplinar).

CONFLICT OF INTEREST STATEMENT

The authors report no conflict of interest.

DATA AVAILABILITY STATEMENT

Data are available upon request.

CLINICAL TRIAL REGISTRATION

Part of this research was registered as a clinical trial at ClinicalTrials.gov (ID: NCT03531346).

ORCID

Izabela K. Garaszczuk https://orcid.org/0000-0002-0560-7069
Maria Romanos-Ibanez https://orcid.org/0009-0007-5221-136X
Alejandra Consejo https://orcid.org/0000-0001-5186-1837

REFERENCES

1. Consejo A, Melcer T, Rozema JJ. Introduction to machine learning for ophthalmologists. *Semin Ophthalmol*. 2019;34:19–41.
2. Murube J. Tear osmolarity. *Ocul Surf*. 2006;4:62–73.
3. Tomlinson A, Khanal S. Assessment of tear film dynamics: quantification approach. *Ocul Surf*. 2005;3:81–95.
4. Potvin R, Makari S, Rapuano CJ. Tear film osmolarity and dry eye disease: a review of the literature. *Clin Ophthalmol*. 2015;9:2039–47.
5. Lemp MA, Bron AJ, Baudouin C, Benítez del Castillo JM, Geffen D, Tauber J, et al. Tear osmolarity in the diagnosis and management of dry eye disease. *Am J Ophthalmol*. 2011;151:792–8.e1.
6. Craig JP, Nichols KK, Akpek EK, Caffery B, Dua HS, Joo CK, et al. TFOS DEWS II definition and classification report. *Ocul Surf*. 2017;15:276–83.
7. Wolffsohn JS, Arita R, Chalmers R, Djalilian A, Dogru M, Dumbleton K, et al. TFOS DEWS II diagnostic methodology report. *Ocul Surf*. 2017;15:539–74.
8. Pritchard N, Fonn D, Brazeau D. Discontinuation of contact lens wear: a survey. *Int Contact Lens Clin*. 1999;26:157–62.
9. Richdale K, Sinnott LT, Skadahl E, Nichols JJ. Frequency of and factors associated with contact lens dissatisfaction and discontinuation. *Cornea*. 2007;26:168–74.
10. Guillon M, Maissa C. Dry eye symptomatology of soft contact lens wearers and nonwearers. *Optom Vis Sci*. 2005;82:829–34.
11. Chalmers RL, Begley CG. Dryness symptoms among an unselected clinical population with and without contact lens wear. *Cont Lens Anterior Eye*. 2006;29:25–30.
12. Nichols KK, Redfern RL, Jacob JT, Nelson JD, Fonn D, Forstot SL, et al. The TFOS International Workshop on Contact Lens Discomfort: report of the definition and classification subcommittee. *Invest Ophthalmol Vis Sci*. 2013;54:TFOS14–FOS19.
13. Iskeleli G, Karakoç Y, Aydin O, Yetik H, Uslu H, Kizilkaya M. Comparison of tear-film osmolarity in different types of contact lenses. *CLAO J*. 2002;28:174–6.
14. Stahl U, Ho A, Brent G, Naduvilath T, Stapleton F. Measurements of solutions and contact lenses with a vapor pressure osmometer. *Optom Vis Sci*. 2007;84:321–7.
15. Ramamoorthy P, Sinnott LT, Nichols JJ. Contact lens material characteristics associated with hydrogel lens dehydration. *Ophthalmic Physiol Opt*. 2010;30:160–6.
16. González-Méijome JM, López-Alemany A, Almeida JB, Parafita MA. Dynamic in vitro dehydration patterns of unworn and worn silicone hydrogel contact lenses. *J Biomed Mater Res B Appl Biomater*. 2009;90:250–8.
17. Garaszczuk IK, Mousavi M, Szczesna-Iskander DH, Cerviño A, Iskander DR. A 12-month prospective study of tear osmolarity in contact lens wearers refitted with daily disposable soft contact lenses. *Optom Vis Sci*. 2020;97:178–85.
18. Tavakoli A, Markoulli M, Flanagan J, Papas E. The validity of point of care tear film osmometers in the diagnosis of dry eye. *Ophthalmic Physiol Opt*. 2022;42:140–8.
19. Koh S, Ikeda C, Watanabe S, Oie Y, Soma T, Watanabe H, et al. Effect of non-invasive tear stability assessment on tear meniscus height. *Acta Ophthalmol*. 2015;93:e135–e139.
20. Nichols KK. Patient-reported symptoms in dry dye disease. *Ocul Surf*. 2006;4:137–45.
21. Tong L, Zhou XY, Jylha A, Aapola U, Liu DN, Koh SK, et al. Quantitation of 47 human tear proteins using high resolution multiple reaction

monitoring (HR-MRM) based-mass spectrometry. *J Proteomics*. 2015;115:36–48.
22. Tomlinson A, Khanal S, Ramaesh K, Diaper C, McFadyen A. Tear film osmolarity: determination of a referent for dry eye diagnosis. *Invest Ophthalmol Vis Sci*. 2006;47:4309–15.
23. Tian L, Qu J, Sun X. Repeatability and reproducibility of noninvasive keratograph 5M measurements in patients with dry eye disease. *J Ophthalmol*. 2016;2016:8013621. https://doi.org/10.1155/2016/8013621
24. Hong J, Sun X, Wei A, Cui X, Li Y, Qian T, et al. Assessment of tear film stability in dry eye with a newly developed keratograph. *Cornea*. 2013;32:716–21.
25. Mousavi M, Jesus DA, Garaszczuk IK, Szczesna-Iskander DH, Iskander DR. The utility of measuring tear film break-up time for prescribing contact lenses. *Cont Lens Anterior Eye*. 2018;41:105–9.
26. Begley C, Caffery B, Chalmers R, Situ P, Simpson T, Nelson JD. Review and analysis of grading scales for ocular surface staining. *Ocul Surf*. 2019;17:208–20.
27. Rico-del-Viejo L, Lorente-Velázquez A, Hernández-Verdejo JL, García-Mata R, Benítez-del-Castillo JM, Madrid-Costa D. The effect of ageing on the ocular surface parameters. *Cont Lens Anterior Eye*. 2018;41:5–12.
28. Garcia-Queiruga J, Pena-Verdeal H, Sabucedo-Villamarin B, Garcia-Resua C, Giraldez MJ, Yebra-Pimentel E. Analysis of the differences in ocular surface damage and inflammatory signs between healthy and evaporative dry eye participants. *Ocul Immunol Inflamm*. 2023;31:970–7.
29. Khanal S, Tomlinson A, Diaper CJ. Tear physiology of aqueous deficiency and evaporative dry eye. *Optom Vis Sci*. 2009;86:1235–40.
30. Chen SP, Massaro-Giordano G, Pistilli M, Schreiber CA, Bunya VY. Tear osmolarity and dry eye symptoms in women using oral contraception and contact lenses. *Cornea*. 2013;32:423–8.
31. Versura P, Campos EC. TearLab® Osmolarity System for diagnosing dry eye. *Expert Rev Mol Diagn*. 2013;13:119–29.
32. Jacobi C, Jacobi A, Kruse FE, Cursiefen C. Tear film osmolarity measurements in dry eye disease using electrical impedance technology. *Cornea*. 2011;30:1289–92.
33. Abdelmotaal H, Hazarbasanov R, Taneri S, Al-Timemy A, Lavric A, Takahashi H, et al. Detecting dry eye from ocular surface videos based on deep learning. *Ocul Surf*. 2023;28:90–8.
34. Arita R, Itoh K, Inoue K, Kuchiba A, Yamaguchi T, Amano S. Contact lens wear is associated with decrease of meibomian glands. *Ophthalmology*. 2009;116:379–84.
35. Alghamdi WM, Markoulli M, Holden BA, Papas EB. Impact of duration of contact lens wear on the structure and function of the meibomian glands. *Ophthalmic Physiol Opt*. 2016;36:120–31.
36. Llorens-Quintana C, Rico-Del-Viejo L, Syga P, Madrid-Costa D, Iskander DR. A novel automated approach for infrared-based assessment of meibomian gland morphology. *Transl Vis Sci Technol*. 2019;8:17. https://doi.org/10.1167/tvst.8.4.17
37. Green-Church KB, Butovich I, Willcox M, Borchman D, Paulsen F, Barabino S, et al. The international workshop on meibomian gland dysfunction: report of the subcommittee on tear film lipids and lipid–protein interactions in health and disease. *Invest Ophthalmol Vis Sci*. 2011;52:1979–93.
38. Pult H, Riede-Pult BH. Non-contact meibography in diagnosis and treatment of non-obvious meibomian gland dysfunction. *J Optom*. 2012;5:2–5.
39. Llorens-Quintana C, Rico-Del-Viejo L, Syga P, Madrid-Costa D, Iskander DR. Meibomian gland morphology: the influence of structural variations on gland function and ocular surface parameters. *Cornea*. 2019;38:1506–12.
40. Llorens-Quintana C, Garaszczuk IK, Szczesna-Iskander DH. Meibomian glands structure in daily disposable soft contact lens wearers: a one-year follow-up study. *Ophthalmic Physiol Opt*. 2020;40:607–16.
41. Pult H, Riede-Pult BH. Non-contact meibography: keep it simple but effective. *Cont Lens Anterior Eye*. 2012;35:77–80.

How to cite this article: Garaszczuk IK, Romanos-Ibanez M, Consejo A. Machine learning-based prediction of tear osmolarity for contact lens practice. *Ophthalmic Physiol Opt*. 2024;00:1–10. https://doi.org/10.1111/opo.13302