Title

Locality is the strongest predictor of expert performance in image-based differentiation of bacterial and fungal corneal ulcers from India.

Journal

Indian Journal of Ophthalmology, 72(4)

Authors

Rosenberg, Christopher
Prajna, Venkatesh
Srinivasan, Muthiah
et al.

Publication Date

2024-04-01

DOI

10.4103/IJO.IJO_3396_22

Peer reviewed

Original Article

Locality is the strongest predictor of expert performance in image-based differentiation of bacterial and fungal corneal ulcers from India

Christopher R Rosenberg, Venkatesh Prajna[1], Muthiah K Srinivasan[1], Prajna C Lalitha[1], Tiru Krishnan[2], Revathi Rajaraman[3], Anitha Venugopal[4], Nisha Acharya[5], Gerami D Seitzman[5], Jennifer Rose-Nussbaumer[6], Maria A Woodward[7], Thomas M Lietman[5], John Peter Campbell, Jeremy D Keenan[5], Travis K Redd
*On behalf of the Corneal Ulcer Image Interpretation Study Group**

Purpose: This study sought to identify the sources of differential performance and misclassification error among local (Indian) and external (non-Indian) corneal specialists in identifying bacterial and fungal keratitis based on corneal photography. **Methods:** This study is a secondary analysis of survey data assessing the ability of corneal specialists to identify acute bacterial versus fungal keratitis by using corneal photography. One-hundred images of 100 eyes from 100 patients with acute bacterial or fungal keratitis in South India were previously presented to an international cohort of cornea specialists for interpretation over the span of April to July 2021. Each expert provided a predicted probability that the ulcer was either bacterial or fungal. Using these data, we performed multivariable linear regression to identify factors predictive of expert performance, accounting for primary practice location and surrogate measures to infer local fungal ulcer prevalence, including locality, latitude, and dew point. In addition, Brier score decomposition was used to determine experts' *reliability* ("calibration") and *resolution* ("boldness") and were compared between local (Indian) and external (non-Indian) experts. **Results:** Sixty-six experts from 16 countries participated. Indian practice location was the only independently significant predictor of performance in multivariable linear regression. Resolution among Indian experts was significantly better (0.08) than among non-Indian experts (0.01; $P < 0.001$), indicating greater confidence in their predictions. There was no significant difference in reliability between the two groups ($P = 0.40$). **Conclusion:** Local cornea experts outperformed their international counterparts independent of regional variability in tropical risk factors for fungal keratitis. This may be explained by regional characteristics of infectious ulcers with which local corneal specialists are familiar.

Access this article online
Website: https://journals.lww.com/ijo
DOI: 10.4103/IJO.IJO_3396_22
Quick Response Code:

Key words: Bacterial keratitis, corneal ulcers, expert prediction, fungal keratitis

Casey Eye Institute, Oregon Health and Science University, Portland, OR, USA, [1]Department of Ophthalmology, Aravind Eye Hospital, Madurai, [2]Department of Ophthalmology, Aravind Eye Hospital, Pondicherry, [3]Department of Ophthalmology, Aravind Eye Hospital, Coimbatore, [4]Department of Ophthalmology, Aravind Eye Hospital, Tirunelveli, Tamil Nadu, India, [5]Francis I. Proctor Foundation, University of California San Francisco, San Francisco, CA, USA, [6]Byers Eye Institute, Stanford University, CA, USA, [7]Kellogg Eye Center, University of Michigan, Michigan, USA

*Study group members: Drs. Diana Alvarez-Melloni, Menen Ayalew, Ashwin Balasubramanian, Elsie Chan, Matilda Chan, Meenu Chaudhary, Thomas Chia, James Chodosh, YY Choong, Joseph Christenbury, Josephine Christy, John Clements, John Dart, Mohammad Dastjerdi, Matthew Denny, Sathish Devarajan, Mohamed Elghobaier, Chris Estopinal, Preethika Gandhi, Nikhil Gokhale, Colleen Halfpenny, Rossen Hazarbassanov, Natalie Hernandez, Anna Hovakimyan, David Hwang, Frank Hwang, Tomas Jaeschke, Vishal Jhanji, Faris Karas, Divya Karthik, Camila Kase, Lakshmi Kattana, Tyson Kim, Aaleya Koreishi, David Liang, Christine Martinez, Rafael Martinez-Costa, Stephen McLeod, Jodhbir S Mehta, Michael Mimouni, Adam Moss, Afshan Nanji, Nathan Nataneli, Vasudha Panday, Sayali Pradhan, Ying Qian, Naveen Rao, Julie Schallhorn, Ruti Sella, Suvitha Selvaraj, David Spokes, Neha Shaik, Nakul Shekhawat, Alan Sugar, Audrey Talley Rostov, Napaporn Tananuvat, Chulaluck Tangmonkongvoragul, Tanya Trinh, Sonal Tuli, Phit Upaphong, Bart van Dooren, Manoj Vasudevan, and Elizabeth Viriya.

Correspondence to: Dr. Travis K Redd, Casey Eye Institute, Oregon Health and Science University, 515 SW Campus Dr, Portland, OR, 97239, USA. E-mail: redd@ohsu.edu

Received: 30-Dec-2022 Revision: 01-Sep-2023
Accepted: 21-Sep-2023 Published: 08-Mar-2024

Infectious keratitis is a major cause of global blindness.[1–3] Directed antimicrobial therapy requires identification of the underlying pathogen, particularly the differentiation between bacterial and fungal keratitis, which together account for >95% of corneal ulcers.[4,5] Cultures of corneal scrapings are the current gold standard for determining the causative pathogen of infectious keratitis but have low sensitivity.[6,7] Even when cultures are positive, the results are typically not available for several days, which may result in a delay in effective therapy and worse visual outcomes. In the absence of microbiologic data, empiric treatment decisions must be made based on human clinical impression.[8–11]

Prior studies indicate that even expert cornea clinicians are only able to correctly differentiate bacterial from fungal

Cite this article as: Rosenberg CR, Prajna V, Srinivasan MK, Lalitha PC, Krishnan T, Rajaraman R, *et al*. Locality is the strongest predictor of expert performance in image-based differentiation of bacterial and fungal corneal ulcers from India. Indian J Ophthalmol 2024;72:526-32.

keratitis approximately two thirds of the time based on clinical impression.[10,12] Our group conducted the largest published survey quantifying expert performance in this task, confirming this relatively low overall human performance but also demonstrating high variability among respondents.[13] In addition, we determined that expert cornea specialists practicing in India significantly outperformed their non-Indian counterparts on this image set of ulcers from South India. There are several potential explanations for this higher level of performance among Indian experts: 1) Greater familiarity with fungal keratitis, which accounts for approximately half of corneal ulcers in India but is rare in more temperate climates;[14] 2) Greater experience with corneal ulcers in general due to the remarkably high burden of disease in this region; and 3) Familiarity with local nuances in the appearance of ulcers from one's own area of practice. To better characterize the underlying reasons for variable performance and inform future educational initiatives to improve the clinical differentiation of bacterial and fungal keratitis, in this study, we evaluated whether overall performance correlated with regional differences in the prevalence of risk factors for fungal keratitis and determined the underlying reasons for misclassification error among these experts.

Methods

Testing set of corneal ulcer images

A tertiary eye care system in South India has conducted several large-scale corneal ulcer trials in recent years.[4,5,15–17] In these trials, many patients underwent corneal photography using a handheld Nikon (Tokyo, Japan) D-series digital single-lens reflex camera according to a standardized lighting and photography protocol. Photographs were repeated until at least one high-quality image was obtained, which was achieved in nearly every case.[4] From these studies, we collated a database of high-quality corneal photographs of approximately 1000 culture-confirmed corneal ulcer cases. All images were obtained from subjects with corneal ulcers at one of four study sites (Madurai 42, Pondicherry 26, Tirunelveli 19, or Coimbatore 13). There were no culture-negative or polymicrobial infections. A discrete testing set of 100 images was randomly selected from this database with stratified random sampling to include 50% bacterial ulcers and 50% fungal ulcers by using the Scikit-learn 0.24.2 package in Python 3.[18]

Expert image graders

Sixty-three cornea specialists were previously recruited, with IRB approval (STUDY00023793), from the home institutions of the authors and via email correspondence through kera-net. The only inclusion criteria for participation were prior completion of a minimum 1-year fellowship in cornea and external disease and access to an internet connection. Residents and current cornea fellows were excluded to avoid underestimating expert human performance. Physicians who recruited or cared for any of the subjects in the above-mentioned randomized trials were excluded to ensure that participants' responses were based only on information presented in the photographs.

Image interpretation

Image reading took place from April to July 2021. Images were presented to participants one at a time in random order with the following prompt: "Based on the image above, what is the likelihood that this ulcer is caused by fungus, not bacteria? Enter a percent value between 0 (absolutely certain it is bacterial) and 100 (absolutely certain it is fungal)." Because bacterial and fungal infections were mutually exclusive and an exhaustive list of the possible causes of infection in this dataset, a respondent's estimated probability of bacterial infection for a given image was equivalent to the complement of the estimated probability of fungal infection [i.e., P (bacterial) =1 – P (fungal)]. All graders were informed that this image set consisted of 50% culture-proven bacterial ulcers and 50% culture-proven fungal infections.[7,14] Graders assessed images sequentially and were unable to revise their predictions. Previous work showed that corneal specialists typically rely on the presence of an irregular/feather border to presume fungal infection, whereas an epithelial plaque or wreath infiltrate more likely indicates bacterial infection.[10] Given these established conclusions, we opted not to include additional questions in our survey to maximize the response rate.

The primary analysis of respondent performance on this image set has been published previously.[13] The principal metric was the area under the receiver operating curve (AUC), which in this case is simply the probability that a randomly selected fungal case would be scored higher (on the fungal probability scale) than a randomly selected bacterial case. Mean performance among cornea specialists practicing primarily in India (where the image set was obtained) demonstrated statistically significantly superior performance (AUC = 0.72) compared to their non-Indian counterparts (AUC = 0.59; $P < 0.001$). Much of this difference appeared to be attributable to superior accuracy in identifying fungal ulcers specifically (76% vs. 49%, $P < 0.001$), whereas no difference was noted in accuracy for identifying bacterial ulcers (71% vs. 71%).

Geographic practice location analysis

We hypothesized that graders practicing in regions with a higher burden of fungal keratitis would demonstrate superior performance on this test set comprising 50% fungal and 50% bacterial ulcers. Tropical environments have been shown to be a key risk factor for fungal keratitis.[19-21] Therefore, we assessed respondents' latitude and dew point (a measure of humidity that accounts for temperature) as surrogate measures of the local climate. Practice location latitudes were acquired from Google Maps (Alphabet Inc., CA, USA). Dew points were computed from World Weather Online (Zoomash Ltd, London, UK) by using monthly mean temperature and relative humidity values for 2019. All statistical analyses were performed in R version 4.0.5.[22]

Misclassification analysis

We used Brier score decomposition to determine and compare the primary sources of misclassification error between Indian and non-Indian respondents.[23] This technique is commonly applied to assess performance in weather prediction and other scenarios where forecasters provide an estimated probability of the occurrence of a categorical outcome, as is the case in this study.[24,25] This involves decomposing the Brier score (effectively the mean squared error of a subject's responses) into its constituent elements: *reliability*, *resolution*, and *uncertainty*. Reliability estimates the "calibration" of a respondent's predictions by measuring the difference between the conditional probability of the outcome (conditioned on the predicted probability) and the predicted probability value itself. Reliability ranges from 0 to 1. For example, among all cases where a respondent predicted a 70% probability of fungal keratitis, the proportion of truly fungal cases would be close to 70% for a respondent with good reliability. Counterintuitively, a lower reliability value results in a lower Brier score and thus indicates better overall performance. In short, a highly capable forecaster will have good internal calibration of their responses resulting in a low reliability value (near 0) and hence a low Brier score.

Resolution (or "boldness") measures how much the conditional probability of an outcome given a particular prediction differs from the overall prevalence of that category in the test set. For example, after being informed that 50% of the ulcers in this dataset are fungal, a respondent with zero resolution (or minimal "boldness") would indicate a 50% probability of fungal etiology for every image graded. This is essentially a measure of the degree to which a respondent is hedging their bets; for example, do they consistently predict at or near 50% probability of fungal keratitis for every image, regardless of its appearance, or are they confident of their predictions and hence more willing to assign them higher probabilities? Resolution ranges from 0 to the level of uncertainty (discussed below), which in this case is 0.5. A higher resolution value results in a lower Brier score and thus indicates better overall performance. In short, a good forecaster will have high confidence in their responses, resulting in a high resolution value (near 0.5).

Uncertainty is the final component of the Brier score decomposition and accounts for the degree of class imbalance, which was nonexistent in this case due to the equal distribution of bacterial and fungal images in the testing set. Thus, uncertainty was not relevant to this analysis. Nonparametric tests (Wilcoxon rank sum) were used to compare resolution and reliability between Indian and non-Indian participants. In this context, the terms "Indian" and "non-Indian" are used to indicate a participant's primary practice location, not their ethnic, racial, or cultural affiliations.

Image-level analysis

To estimate overall human performance on each image, we determined the "ensemble estimated probability," defined as the mean of the predicted probabilities across all respondents. The ensemble estimated probability was compared against the ground truth (whether the image was bacterial or fungal) to identify the test cases with the best overall grader performance (smallest difference between ensemble estimated probability and ground truth) and worst overall grader performance. We also identified the image with the greatest variability in grader responses by measuring the standard deviation of the predictions.

Results

A total of 66 cornea specialists from 16 countries participated in this study [Table 1]. The majority practiced primarily in either India (12 respondents [18%]) or the United States (33 respondents [50%]). The remainder (21 respondents [32%]) were from Argentina, Armenia, Australia, Brazil, Canada, Egypt, Israel, Malaysia, Mexico, Netherlands, Singapore, Spain, Thailand, and the United Kingdom.

The mean AUC, representing grader performance, for the Indian (local) and non-Indian (external) groups was 0.72 and 0.59, respectively. Dew point ranged from 3°C to 25°C (mean: 14°C), and latitude ranged from 1.4° to 52.6° (mean: 31.1°).

Bivariable analyses showed significant positive associations between AUC and Indian practice location ($P < 0.001$) and dew point ($P = 0.003$), and a negative association with absolute latitude ($P = 0.002$) [Fig. 1]. However, only Indian practice location remained an independently significant predictor of grader performance in the multivariable linear regression model accounting for all these covariates ($P = 0.006$) [Table 2].

Brier score decomposition

The overall Brier score among all respondents was 0.56, with a resolution of 0.01 and a reliability of 0.09. Regional comparison of the Brier score decomposition revealed that resolution ("boldness") values were significantly better among Indian graders (0.08) compared to non-Indian graders (0.01; $P < 0.001$ [Wilcoxon Rank Sum test]) [Fig. 2]. The reliability ("calibration") among Indian graders (0.05) appeared slightly better than that among non-Indian graders (0.09), though this difference was not statistically significant ($P = 0.40$).

Image-level analysis

Fig. 3 depicts the images with the "best grader performance," "worst grader performance," and "most variability in grader responses," as well as the corresponding distributions of estimated probabilities provided by respondents.

Discussion

This study quantitatively evaluated the impact of regionality on the ability of expert cornea specialists to perform image-based differentiation of bacterial and fungal keratitis and measured the underlying causes of misclassification error by using the Brier score decomposition. Overall, local experts from the region where the images were obtained (India) significantly outperformed their external (non-Indian) counterparts, independent of surrogate measures of regional variability in fungal keratitis prevalence. Brier score decomposition demonstrated that this difference was likely attributable to superior resolution (or "boldness") among Indian respondents. These results highlight the impact of regional variability in ulcer epidemiology on the ability of even expert cornea specialists to reliably interpret ulcer appearance and inform potential educational initiatives to improve human performance in this task.

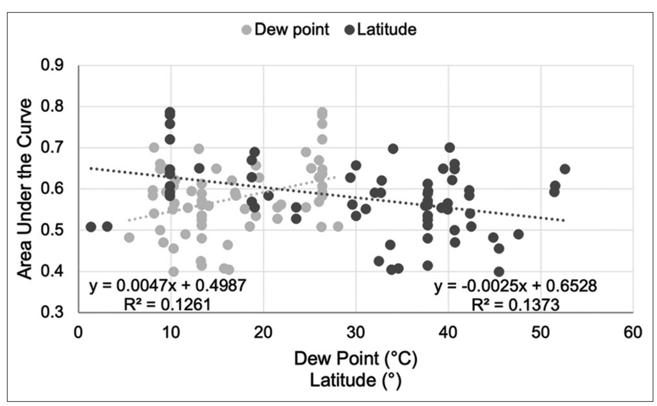

Figure 1: Linear regression of latitude and dew point versus AUC. A lower latitude indicates closer proximity to the equator, thus indicating a tropical location. A higher dew point indicates a greater tropical climate

Table 1: Grader Practice Locations. Regions defined according to absolute value of latitude: Tropical <23.49°; Subtropical 23.50°–45.00°; Temperate >45.00°

Temperate	Count	Subtropical	Count	Tropical	Count
San Francisco, CA	10	Cairo, Egypt	2	Madurai, India	9
New York, NY	3	Arlington, TX	1	Chiang Mai, Thailand	3
Ann Arbor, MI	2	Atlanta, GA	1	Mumbai, India	2
Portland, OR	2	Buenos Aires, Argentina	1	Sao Paulo, Brazil	2
Baltimore, MD	1	Gainesville, FL	1	Chennai, India	1
Boston, MA	1	Haifa, Israel	1	Petaling Jaya, Malaysia	1
Breda, Netherlands	1	Loma Linda, CA	1	Queretaro City, Mexico	1
Burlington, TX	1	San Antonio, TX	1	Singapore, Singapore	1
Columbus, OH	1	Sydney, Australia	1		
Edina, MN	1	Tel Aviv, Israel	1		
London, UK	1	Temple, TX	1		
Melbourne, Australia	1	Tijuana, Mexico	1		
Newark, NJ	1				
Norwich, UK	1				
Philadelphia, PA	1				
Pittsburgh, PA	1				
Richmond, VA	1				
Seattle, WA	1				
Valencia, Spain	1				
Yerevan, Armenia	1				
Total	33		13		20

Table 2: Multivariable Linear Regression of Regional Factors Predicting Grader AUC in Differentiating Bacterial and Fungal Keratitis

Independent variable	Coefficient	*P*
Indian practice location	0.107	0.006
Dew point	−0.001	0.766
Latitude	−0.001	0.736

Local Indian graders demonstrated particularly higher accuracy in identifying fungal keratitis compared to their non-Indian counterparts, prompting the supposition that the regional prevalence of fungal keratitis may be associated with improved performance in this dataset. However, although multiple tropical variables were associated with higher grader performance in bivariable analyses, Indian practice location remained the only independently significant predictor of performance in the multivariable model. It may be that familiarity with the appearance of ulcers in one's own region resulted in the greater performance demonstrated in this study. Alternatively, these Indian experts may simply be more confident and accurate in identifying corneal infections due to their remarkably high local incidence, which is among the largest in the world.[19,23] Evaluating Indian grader performance on non-Indian ulcers will be required to address these questions. Nonetheless, these results suggest that greater clinical experience evaluating corneal ulcers from one's own region may be associated with greater confidence and accuracy in predicting the etiology of infection. This indicates that educational initiatives designed to increase learner exposure to corneal infections may improve performance. Efforts are underway to incorporate the image set from this study into an educational platform to pursue this goal.

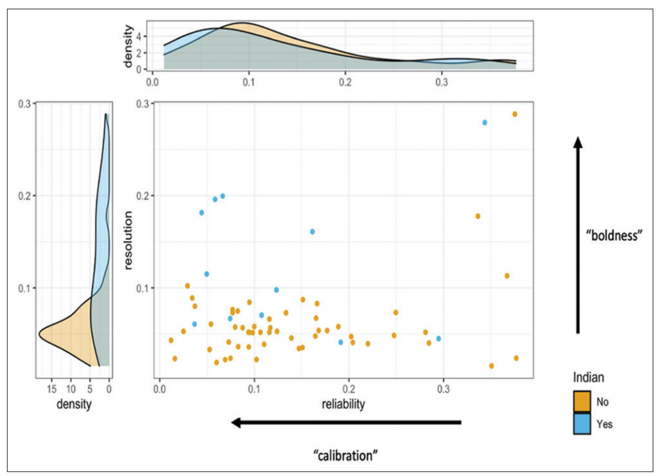

Figure 2: Brier decomposition scatter plot of each respondent. Marginal distributions compare the resolution ($P < 0.001$) and reliability ($P = 0.4$) of cornea specialists for identifying the cause of corneal infection

The overall accuracy of mean responses among all experts surveyed in the primary study was 61%, which is similar to the results of previous smaller studies.[10,12] The consistency of these accuracy measurements across studies establishes that even many expert humans are not reliably able to distinguish between bacterial and fungal keratitis based on appearance alone. This raises the question of whether the clinical appearance of an ulcer does not contain sufficient information to reliably determine the underlying etiology of infection, in

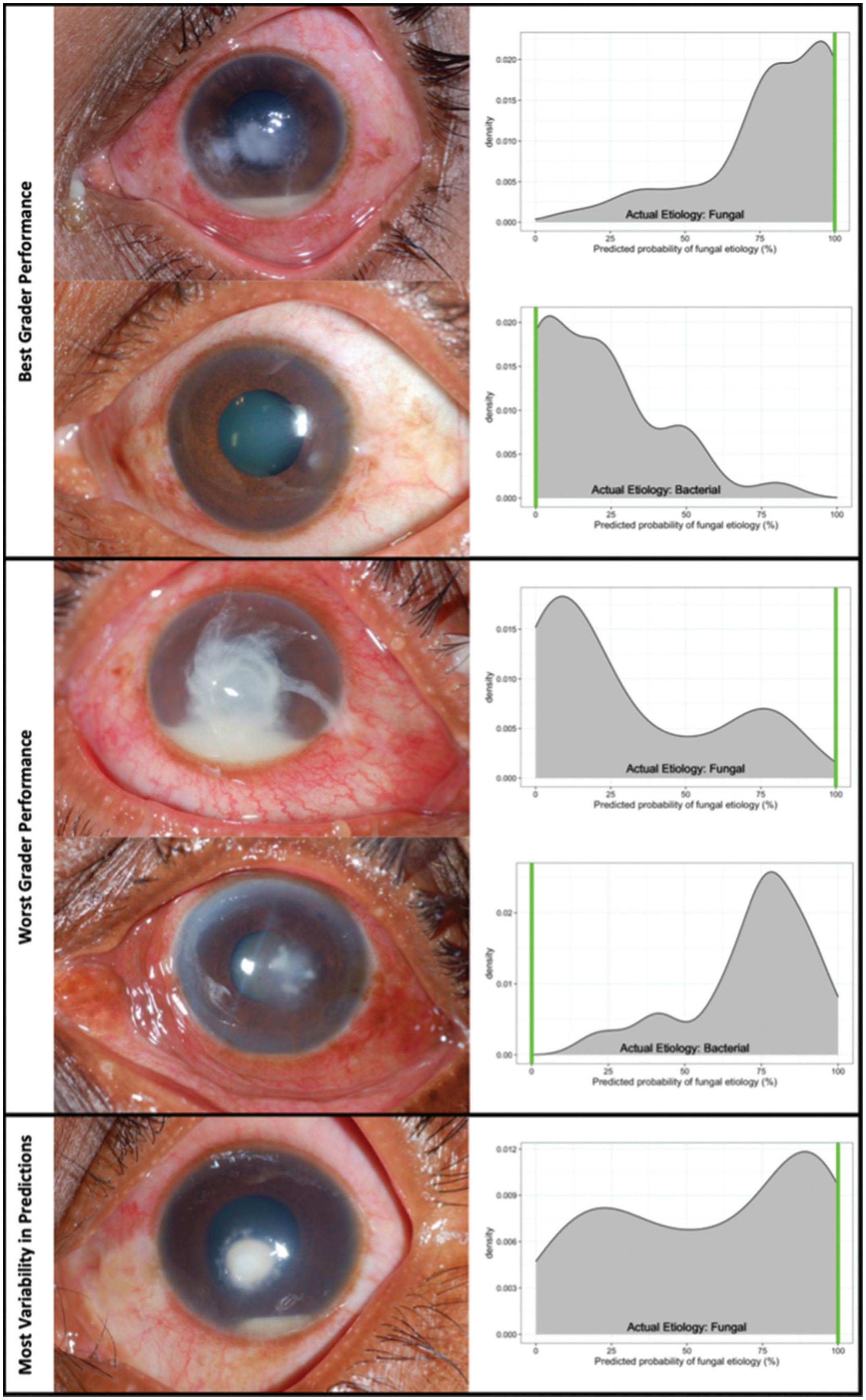

Figure 3: Images for best, worst, and most variability in grader predictions. Corresponding distributions of estimated probabilities of fungal etiology are illustrated

which case humans may be approximating Bayes error (the best possible performance given the information provided). However, computer vision models have achieved superior performance in analyzing these same images, indicating that there is valuable predictive information being missed by many human experts in interpreting ulcer appearance to determine the underlying cause of infection.[26] Additional interrogation of how artificial intelligence models extract this information may inform educational initiatives to improve human performance in this task.

Several potential limitations of this study must be considered. First, it is possible that the variables used to estimate fungal keratitis burden, that is, country, latitude, and dew point, do not in every case closely correlate with actual disease prevalence. For example, it is possible that Indian cornea specialists see such a high volume of infectious keratitis, even compared to specialists in other tropical countries, that the tropical variables used in this study could not adequately capture this disproportion. Institutional pedigrees were not obtained from the respondents. Therefore, we are unable to assess if training institutions had an impact on grader performance. A future comparison between Indian specialists who trained or practiced at different institutions or regions could help answer this question. Agricultural employment is also a known risk factor for fungal keratitis because of the propensity for vegetative trauma.[19,20] Unfortunately, standardized data representing the agricultural workforce are not accessible; thus, this variable could not be included in our model. Therefore, geographic location and climate remain the best surrogate measures of regional fungal ulcer prevalence given the scarcity of other relevant epidemiologic data.

Image quality alone can have a significant impact on human grader performance. However, in this study, all images were obtained by professional photographers using a standardized protocol and were confirmed to be of adequate quality for interpretation by a cornea specialist (TR) prior to subject recruitment. Furthermore, this dataset does not include culture-negative, polymicrobial, or non-bacterial/non-fungal infections. The exclusion of culture-negative cases was necessary to ensure an accurate gold-standard label for all images, and we elected to narrow the focus to bacterial and fungal ulcers because together they account for 95% of ulcer cases.[14] Data collection efforts are ongoing to establish a database of all ulcers presenting to a tertiary eye care system in South India, including culture-negative and non-bacterial/non-fungal infections, to allow a more granular assessment of human performance in recognizing all types of infectious etiologies. Finally, this is a secondary analysis and was not pre-specified; thus, these results should be interpreted primarily as hypothesis generation rather than establishment of causality.

Conclusion

Cornea specialists who practice in India showed greater ability to differentiate fungal and bacterial keratitis in this image-based test, largely due to having greater confidence in their predictions and perhaps greater familiarity with local corneal ulcer phenotypes. These results suggest that educational initiatives designed to increase clinicians' exposure to different types of corneal ulcers may help close this performance gap in clinical determination of the cause of infection, which may enable more rapid initiation of effective antimicrobial therapy and improved visual outcomes.

Acknowledgements

The authors would like to thank Bill Stanford, database administrator for the Casey Reading Center, for his help in developing and maintaining the web-based image grading interface for this study.

Financial support and sponsorship

This study was supported by the National Institutes of Health (NIH K12EY027720 [TKR], core grant P30EY10572 [TKR, JPC], and R01EY031033 [MAW]), unrestricted departmental funding provided by Research to Prevent Blindness, and a Research to Prevent Blindness Career Advancement Award (MAW). The funding organization had no role in the design or conduct of this research.

Conflicts of interest

There are no conflicts of interest.

This study involves human participants and was approved by the Oregon Health and Science University Institutional Review Board, ID: STUDY00022423.

References

1. Flaxman SR, Bourne RR, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, *et al*. Global causes of blindness and distance vision impairment 1990–2020: A systematic review and meta-analysis. Lancet Glob Health 2017;5:e1221-34.
2. Whitcher JP, Srinivasan M, Upadhyay MP. Corneal blindness: A global perspective. Bull World Health Organ 2001;79:214-21.
3. Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. Br J Ophthalmol 2012;96:614-8.
4. Srinivasan M, Mascarenhas J, Rajaraman R, Ravindran M, Lalitha P, Glidden DV, *et al*. Corticosteroids for bacterial keratitis: The steroids for corneal ulcers trial (SCUT). Arch Ophthalmol 2012;130:143-50.
5. Prajna NV, Krishnan T, Mascarenhas J, Rajaraman R, Prajna L, Srinivasan M, *et al*. The mycotic ulcer treatment trial: A randomized trial comparing natamycin vs voriconazole. JAMA Ophthalmol 2013;131:422-9.
6. Mcleod SD, Kolahdouz-isfahani A, Rostamian K, Flowers CW, Lee PP, Mcdonnell P. The role of smears, cultures, and antibiotic sensitivity testing in the management of suspected infectious keratitis. Ophthalmology 1996;103:23-8.
7. Varaprasathan G, Miller K, Lietman T, Whitcher JP, Cevallos V, Okumoto M, *et al*. Trends in the etiology of infectious corneal ulcers at the F. I. Proctor Foundation. Cornea 2004;23:360-4.
8. Thomas PA. Fungal infections of the cornea. Eye 2003;17:852-62.
9. Rosa RH, Miller D, Alfonso EC. The changing spectrum of fungal keratitis in South Florida. Ophthalmology 1994;101:1005-13.
10. Dalmon C, Porco TC, Lietman TM, Prajna NV, Prajna L, Das MR, *et al*. The clinical differentiation of bacterial and fungal keratitis: A photographic survey. Invest Ophthalmol Vis Sci 2012;53:1787-91.
11. Thomas PA, Leck AK, Myatt M. Characteristic clinical features as an aid to the diagnosis of suppurative keratitis caused by filamentous fungi. Br J Ophthalmol 2005;89:1554-8.
12. Dahlgren MA, Lingappan A, Wilhelmus KR. The clinical diagnosis of microbial keratitis. Am J Ophthalmol 2007;143:940-4.
13. Redd TK, Prajna NV, Srinivasan M, Lalitha P, Krishnan T, Rajaraman R, *et al*. Expert performance in visual differentiation of bacterial and fungal keratitis. Ophthalmology 2022;129:227-30.

14. Srinivasan M, Gonzales CA, George C, Cevallos V, Mascarenhas JM, Asokan B, *et al*. Epidemiology and aetiological diagnosis of corneal ulceration in Madurai, South India. Br J Ophthalmol 1997;81:965-71.
15. Prajna NV, Krishnan T, Rajaraman R, Patel S, Srinivasan M, Das M, *et al*. Effect of oral voriconazole on fungal keratitis in the mycotic ulcer treatment trial II (MUTT II): A randomized clinical trial. JAMA Ophthalmol 2016;134:1365-72.
16. Narayana S, Krishnan T, Ramakrishnan S, Samantaray PP, Austin A, Pickel J, *et al*. Mycotic antimicrobial localized injection: A randomized clinical trial evaluating intrastromal injection of voriconazole. Ophthalmology 2019;126:1084-9.
17. Prajna NV, Radhakrishnan N, Lalitha P, Austin A, Ray KJ, Keenan JD, *et al*. Cross-linking–assisted infection reduction: A randomized clinical trial evaluating the effect of adjuvant cross-linking on outcomes in fungal keratitis. Ophthalmology 2020;127:159-66.
18. Pedregosa A. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12:2825-30.
19. Brown L, Leck AK, Gichangi M, Burton MJ, Denning DW. The global incidence and diagnosis of fungal keratitis. Lancet Infect Dis 2021;21:e49-57.
20. Deorukhkarl S, Katiyar R, Saini S. Epidemiological features and laboratory results of bacterial and fungal keratitis: A five-year study at a rural tertiary-care hospital in western Maharashtra, India. Singapore Med J 2012;53:264-7.
21. Bharathi MJ, Ramakrishnan R, Meenakshi R, Padmavathy S, Shivakumar C, Srinivasan M. Microbial keratitis in South India: Influence of risk factors, climate, and geographical variation. Ophthalmic Epidemiol 2007;14:61-9.
22. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
23. Murphy AH. General decompositions of MSE-based skill scores: Measures of some basic aspects of forecast quality. Mon Weather Rev 1996;124:2353-69.
24. Hernández-Orallo J, Flach P, Ferri C. A unified view of performance metrics: Translating threshold choice into expected classification loss. J Mach Learn Res 2012;13:2813-69.
25. Ferro CAT, Fricker TE. A bias-corrected decomposition of the Brier score. Q J R Meteorol Soc 2012;138:1954-1960.
26. Redd TK, Prajna NV, Srinivasan M, Lalitha P, Krishnan T, Rajaraman R, *et al*. Image-based differentiation of bacterial and fungal keratitis using deep convolutional neural networks. Ophthalmol Sci 2022;2:100119.