Detection of subclinical keratoconus using machine learning algorithms

[1] Basina swathi , [2] Dr. P. Rajesh kumar

[1,2] Department of Electronics and Communication Engineering

Andhra University College of Engineering , Visakhapatnam

Abstract: A non-inflammatory corneal condition called keratoconus frequently affects both eyes. Due to corneal deformation and scarring, the bilateral ecstatic illness known as keratoconus can impair vision. The prevalence of keratoconus varies from one in 375 individuals in geographical area to as high as one in forty eight in varied ethnic teams. Studies indicate the next incidence and quicker advancement in Middle-Eastern, American, and Asian populations. The disease normally starts when pubescence and progresses over the following 2 to a few decades at a varied rate. because the condition worsens, corneal distortion could become thus severe that patients can not be ready to see to a tolerable degree to perform while not the utilization of soft or hard contact lenses. Contact lenses aren't invariably well-tolerated, and vision loss will considerably lower quality of life... Keratoconus could be a disorder defined by progressive thinning and distortion of the cornea. If detected at an early stage, corneal scleroprotein cross-linking can stop unwellness progression and additional visual loss. though advanced forms ar simply detected, reliable identification of subclinical unwellness are problematic. many totally different machine learning algorithms are wont to improve the detection of subclinical astigmia supported the analysis of multiple forms of clinical measures, like corneal imaging, aberrometry, or biomechanical measurements. The aim of this study is to develop and check a machine learning rule which is able to detect keratoconus at early stages.The projected model would possibly aid physicians in assessing corneal standing and detection keratoconus, that's otherwise difficult through subjective evaluations, significantly at the diagnosing and early stages of the disease.

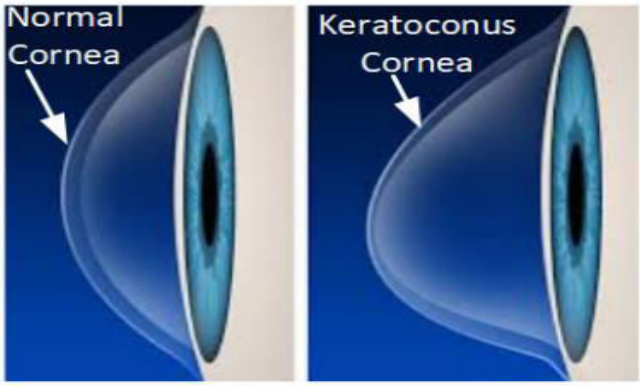
1. Introduction

Since the development of corneal collagen cross-linking, the need for earlier keratoconus identification has increased (CXL). Following the administration of riboflavin (vitamin B2), the cornea is subjected to a photochemical procedure using UV-A light that can stop the progression of keratoconus in 98.3% of instances, even in quite advanced cases. Early diagnosis by community-based optometrists is difficult to improve since asymptomatic individuals with preclinical disease are unlikely to seek assessment. However, the advantage of early therapy to limit visual loss is evident, and there is evidence that it is cost-effective. It's likely that better access or effective community screening using pricy imaging technology will be needed for improved detection.

An ophthalmologist or an optometrist will typically diagnose keratoconus through an eye exam, along with other tests including going over the patient's medical and family history.

The shape of the patient's cornea is frequently determined using a variety of instruments and technology.

Eye refraction: within which the doctor can use special instrumentation that measures eyes to examine for vision issues.



Slit-lamp examination : within which the doctor can measure the form of the cornea and appears for alternative potential issues within the patient's eye, by directional a vertical beam of sunshine on its surface and employing a powerless magnifier to look at it.

Keratometry: during this take a look at the attention doctor focuses a circle of sunshine on the cornea and measures the

reflection to see the fundamental form of the cornea.

Computerized corneal mapping: that consists of special photographic tests, like corneal tomography and corneal topography. it's essentially recording pictures to make an in depth form map of the cornea. ..

Corneal Topography: by taking a picture of the eye with an instrument called corneal topographer, we can easily detect the changes of the cone that is developping in the surface of the eye. Also, by using Marco OPD III, which is a corneal topographer that measures the corneal shape, outline map, size... In other words it gives information about the topography of the eye (**shape maps**). It also measures the corneal aberrometry map that measures specific distortions with the eye's structure, which is known as high order aberrations that can be detected in early stages of keratoconus.

Corneal Tomography: By using an OCT machine (Optical Coherence Tomograph) , which measures the corneal thickness and shows the most thinnest parts of it, where the rate of Keratoconus is most progressive. In other words, it is getting many B-scan ultrasounds of the eye 9. It helps to see how thin the cornea is at different locations, especially on the back surface of the cornea which is the first sign of Keratoconus that can be detected. This can be detected well before having to detect any cone shape on the surface of the eye by the slit-lamp exam. This method is called **Global Pachymetry**.

OCT images

An Oct -Optical coherence tomography- imaging may be a noncontact imaging of the cornea with micron-level high resolution. it's used to get pictures of the corneal thickness of normal and keratoconic eyes and alternative diseases.

It can image the deeper layers or the retina beneath the surface where the early stage of the disease starts, that often causes a mucular hole in the retina. It also measures the corneal thickness and shows the most thinnest parts of it, where the rate of Keratoconus is most progressive.

Notice that the thickness measures of the cornea are between 600 and 500, and that is considered normal, but if the measures get below 400 or 300, that is considered an anomaly and there is a strong probability that the eye is infected with keratoconus disease

1.2 Input Data Types

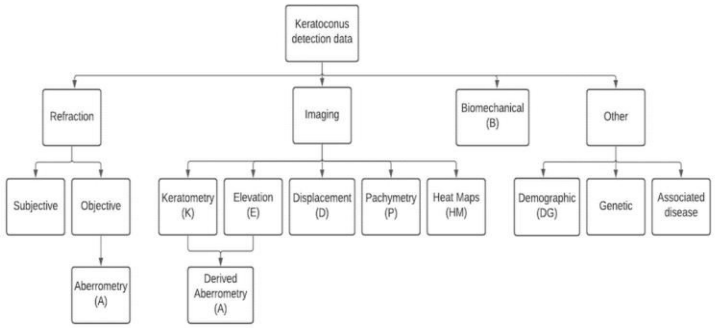


Fig.1. 3(a) : Organizational diagram of relevant data types used for the detection of subclinical keratoconus.

2. Literature survey

To help clinicians have enough time to choose the best course of treatment, Hallett et al. suggested a deep learning-based unsupervised and semi-supervised classification model to notice keratoconus early [4]. Utilizing 124 keratoconus eyes, they earned an accuracy level of 80.3%. Their small sample size, meanwhile, would possibly build it difficult to generalise their results. A supply regression applied math model was applied in [5] to identify cases of keratoconus in its early stages. However, the auto-keratometer was the only corneal used in employed in this investigation. using cornea form knowledge from OCT-based devices, the authors of [6] instructed a classification technique and achieved an accuracy level of ninety two with 244 eyes. However, it's unknown however severe the keratoconus eyes are and whether or not they are affected.In order to guide the implantation of intracorneal ring segments, machine learning has also been employed in the medical aid of keratoconus [7]. may be} encouraging since it demonstrates however AI models can be used to improve supply in varied keratoconus management areas. Recently, a quick summary of varied machine learning ways for keratoconus detection was provided [8]. to boot, current analysis has emphatic the perform and significance of developing computing (AI) systems for the hindrance and watching of keratoconus [9]. though AI models have to this point created encouraging results, extra work is required to encourage the event of a lot of precise algorithms for identifying keratoconus, particularly at its earliest phases.NWe give a thorough outline of previous analysis on key machine learning models, together with multi-layer perceptrons, support vector machines (SVM), unsupervised machine learning algorithms, artificial neural networks, radial basis networks, convolutional neural networks, and decision tree techniques that are developed to notice keratoconus, within the results presented below (Figure 2). using large scale

multi-center datasets gathered from varied membrane clinics, the algorithmic rule was developed with the goal of developing and corroboratory machine learning algorithms that accurately determine early stage keratoconus There are not several analysis that have used totally different machine learning ways on the same collection of knowledge. On an equivalent knowledge set of thirty-nine traditional management eyes and forty nine eyes with subclinical keratoconus, Cao et colleagues examined eight machine learning techniques. The Pentacam picturing was used in conjunction with age, sex, and 9 membrane characteristics, and therefore the authors discovered that random forest, SVM, and K-nearest neighbours had the most effective performance. With an terrorist group of zero.97, random forests had the best sensitivity (94%), SVM had the best specificity (90%), and K-nearest neighbours had the best terrorist group. They used 10-fold cross-validation to substantiate their findings, though it might still be helpful to redo the analysis on an even bigger external knowledge set. Ambrosio et al. divided the classification into four categories: normal, keratoconus, extremely asymmetrical ectasia-ectatic, and subclinical keratoconus. They did this study exploitation techniques like as supply regression, SVMs, and random forests (very uneven ectasia-normal). They enclosed 480 traditional eyes, 204 eyes with keratoconus, seventy two eyes categorized as extraordinarily asymmetrical ectasia-ectatic, and ninety four eyes with subclinical astigmatism. They used each Scheimpflug CT and biomechanical knowledge. The random forest model fared the most effective once subclinical astigmatism was taken into consideration, with 90.4% sensitivity and ninety six specificity. the ultimate model, that was referred to as the picturing and Biomechanical Index, was valid exploitation leave-one-out cross-validation, yielding 850 models total. once making an attempt to categorize three sets of eyes, Lopes et al. additionally conducted a comparative analysis and discovered that random forests performed higher (including subclinical eyes). the best comparative analysis for distinctive subclinical astigmatism was performed by Lavric et al. The authors' study enclosed 1970 traditional eyes, 390 astigmatism eyes, and 791 subclinical (FFKC) astigmatism eyes. The analysis, that utilized twenty five totally different machine learning ways, utilized keratometric, pachymetric, and aberrometric knowledge from the CASIA AS-OCT system. SVM, that reached eighty nine.5% sensitivity for classifying the 3 teams at the same time, proven to be the foremost correct algorithmic rule. the employment of the CASIA distention screening index (ESI) for the assessment of astigmatism severity, which can not accord with clinical diagnosing, and therefore the studied parameters' shut ties to the CASIA device, that restricts generalizability to different systems, area unit among the study's drawbacks.

3.Proposed Methodology

There are variety of limitations to this study that will rather be self-addressed in follow-up studies. compared the clump outcome with Casia ESI index and showed that there is a good agreement between our finding and ESI index spectrum . However, to assess the generalizability of this unattended clump approach methodology, it should be valid by different keratoconus indices like Bellin-Ambrosio (BA) index. Therefore, it's required to conduct another study to substantiate but this approach is generalizable to corneal parameters generated by Pentacam instrument by accessing such datasets. Also, the accuracy of this approach are typically valid if the clinical diagnosing labels of all eyes were procurable. However, accessing clinical diagnosing labels for all eyes in such large datasets might be a tough and tedious task. even so, it's useful to assess the planned approach throughout a follow-up study with a dataset that has clinical diagnosing labelsIn this study a large dataset containing 423 features has been used to train the machine learning algorithms. Out of these features, features with highest discriminating power are found out by performing a univariate selection, which is one of the feature extraction techniques.

Feature Selection

When we encounter a dataset that has alot of features it can be difficult to identify the relevant ones that are highly related to the label/target column and can improve the model's accuracy and the irrelevant ones that can impact negatively the model's performance. Thus we need to apply Feature Selection to help us remove, manually or automatically, the irrelevant features in order to improve the model's performance and reduce training time as well as overfitting.

Overfitting: is when a model learns from the training data too well, meaning that it learns the noise and the details to the point that it can no longer generalize and predict new data. There are two important techniques to limit overfitting : 1. Resampling (example: k-fold cross validation) and 2. evaluate the model using a validation dataset.

In simpler words, overfitting is basically : good performance score on the training data and poor performance score on a new data.

model performance: It is the ability of the model to perform a certain task (example: classification) accurately not only with the training data but also when the model is deployed through a website or an app, to perform on a real-time data.

Feature scores obtained using Univariate feature selection

method to find out the best features.

	Specs	score
1	ESI.Posterior.	1730.581132
2	K.Max..8mm..1	1721.651993
3	Score.Posterior.	1650.762822
4	K.Max..10mm..1	1642.335430
5	Steepest.1	1517.656623
6	Steepest.Posterior.	1517.656623
7	Steepest.3	1517.656623
8	DSI.5mm..1	1338.253269
9	SD_P.4mm..1	1337.386853
10	DSI.6mm..1	1337.354302

4. Results and Evolution Metrices

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Confusion Matrix

FORMULAS

$$ACCURACY = \frac{TP + TN}{TP + FP + FN + TN}$$

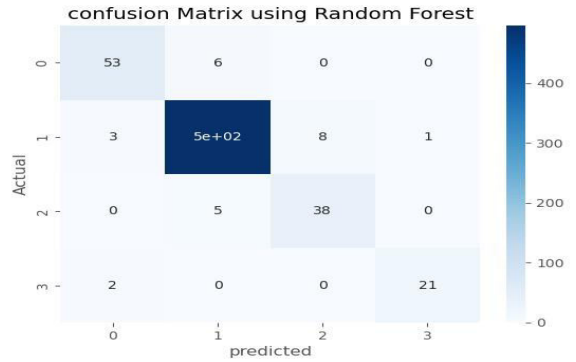
$$PRESICION = \frac{TP}{TP + FP}$$

$$RECALL = \frac{TP}{TP + FN}$$

$$F1 - SCORE = 2 * \frac{Precision * recall}{precision + recall}$$

<u>Predicting keratoconus classes based on the original training data using Random Forest Classifier:</u>

<u>Random Forest</u>

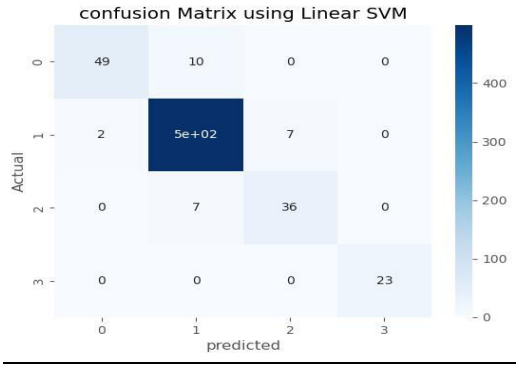


Confusion matrix rf
[[53 6 0 0]
[4 495 8 1]
[0 5 38 0]
[2 0 0 21]]
Accuracy rf 0.9589257503949447
Recall rf 0.9173697355145221
Precision rf 0.9142995913445435

	precision	recall	f1-score	support
1	0.90	0.90	0.90	59
2	0.98	0.97	0.98	508
3	0.83	0.88	0.85	43
4	0.95	0.91	0.93	23
accuracy			0.96	633
macro avg	0.91	0.92	0.92	633
weighted avg	0.96	0.96	0.96	633

<u>Support Vector Machine:</u>

LINEAR SVM

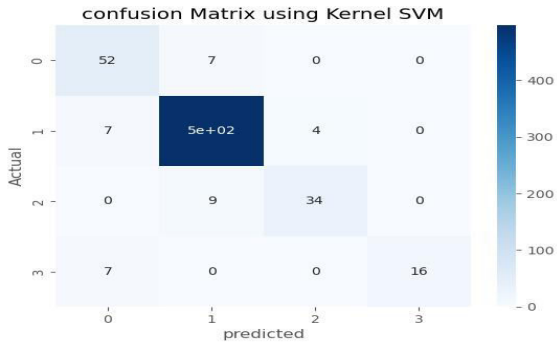


[[49 10 0 0]
[2 499 7 0]
[0 7 36 0]
[0 0 0 23]]
Accuracy linear SVM 0.959

Precision linear SVM 0.941
Recall linear SVM 0.913

	precision	recall	f1-score	support
1	0.96	0.83	0.89	59
2	0.97	0.98	0.97	508
3	0.84	0.84	0.84	43
4	1.00	1.00	1.00	23
accuracy			0.96	633
macro avg	0.94	0.91	0.93	633
weighted avg	0.96	0.96	0.96	633

KERNEL SVM

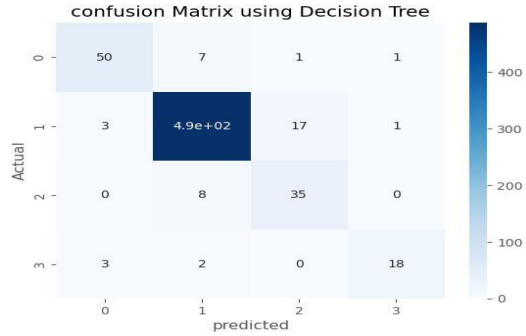


[[52 7 0 0]
[7 497 4 0]
[0 9 34 0]
[7 0 0 16]]

Accuracy rbf SVM 0.946
Precision rbf SVM 0.913
Recall rbf SVM 0.837

	precision	recall	f1-score	support
1	0.79	0.88	0.83	59
2	0.97	0.98	0.97	508
3	0.89	0.79	0.84	43
4	1.00	0.70	0.82	23
accuracy			0.95	633
macro avg	0.91	0.84	0.87	633
weighted avg	0.95	0.95	0.95	633

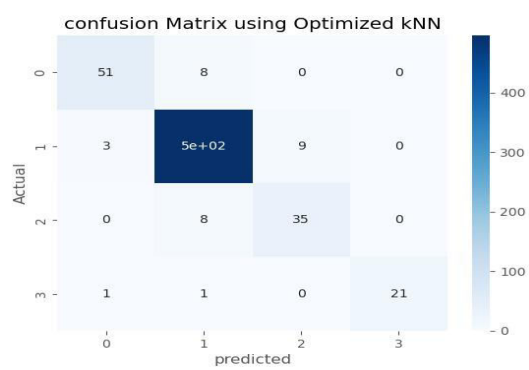
Decision Tree



Confusion matrix dt
[[50 7 1 1]
[4 486 16 2]
[0 6 37 0]
[3 2 0 18]]
Accuracy dt 0.933649289099526
Recall dt 0.8618060881089287
Precision dt 0.8473952262559259

	precision	recall	f1-score	support
1	0.88	0.85	0.86	59
2	0.97	0.96	0.96	508
3	0.69	0.86	0.76	43
4	0.86	0.78	0.82	23
accuracy			0.93	633
macro avg	0.85	0.86	0.85	633
weighted avg	0.94	0.93	0.94	633

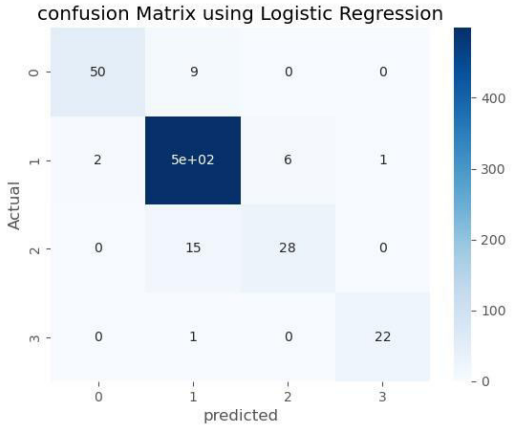
Optimized K-Nearest Neighbors:



Confusion matrix knn
[[51 8 0 0]
[3 495 10 0]
[0 7 36 0]
[1 1 0 21]]
Accuracy knn 0.95260663507109

	precision	recall	f1-score	support
1	0.93	0.86	0.89	59
2	0.97	0.97	0.97	508
3	0.78	0.84	0.81	43
4	1.00	0.91	0.95	23
accuracy			0.95	633
macro avg	0.92	0.90	0.91	633
weighted avg	0.95	0.95	0.95	633

Logistic Regression




```
array([[ 50,   9,   0,   0],
       [  2, 499,   6,   1],
       [  0,  15,  28,   0],
       [  0,   1,   0,  22]])
```

Accuracy: 0.9462875197472354
Precision: 0.9234699221923699
Recall: 0.8593564053784206

Summary:

Algorithm	Accuracy	Recall	Precision
Optimized kNN	0.952	0.891	0.922
Decision Tree	0.933	0.861	0.847
Random Forest	0.958	0.917	0.914
linear SVM	0.946	0.913	0.941
kernel SVM	0.946	0.837	0.913
Logistic Regression	0.946	0.859	0.923

In this project, I have used a dataset containing 3162 eyes of patients and a large number of parameters (423) from which only few were selected due to their feature importance using univariate selection method as they've found to have highest discriminating power in classifying patients with keratoconus.
Out of the six machine learning algorithms used Random forest classifier has the highest accuracy of 95.8% whereas Decision tree algorithm has found to be least accurate of them all.
When comparing Recall scores Random forest algorithm has performed well of the six algorithms tested with a score of 91.7% and kernel SVM has underperformed of all the six.
Linear SVM has highest precision score of all the algorithms tested with a precision score of 94.1% and Decision has found to be least precise with a precision score of 84.7%
So, in my observation Random forest classifier has performed well in classifying the healthy versus keratoconus eyes.

5.Conclusion

Using automated unsupervised machine learning algorithms with corneal topography, tomographic, and thickness profiles, keratoconus status and severity may now be reliably determined.This strategy can be applied in corneal clinics and research settings to enhance our knowledge of corneal alterations in keratoconus and to better diagnose, track changes and progression

6.References

1 : Yousefi S, Yousefi E, Takahashi H, Hayashi T, Tampo H, Inoda S, et al. (2018) Keratoconus severity identification using unsupervised machine learning. PLoS ONE 13(11): e0205998. https://doi.org/10.1371/journal. pone.0205998

2. Davidson AE, Hayes S, Hardcastle AJ, Tuft SJ. The pathogenesis of keratoconus. Eye (Lond) 2014 Feb;28(2):189-195 [doi: 10.1038/eye.2013.278] [Medline: 24357835]

3. Godefrooij DA, de Wit GA, Uiterwaal CS, Imhof SM, Wisse RP. Age-specific incidence and prevalence of keratoconus: a nationwide registration study. Am J Ophthalmol 2017 Mar;175:169-172. [doi25 10.1016/j.ajo.2016.12.015] [Medline: 28039037]

4.N. Hallett, K. Yi, J. Dick, C. Hodge, G. Sutton, Y. G. Wang, and

J. You, ''Deep learning based unsupervised and semi-supervised classification for keratoconus,'' Jan. 2020, arXiv:2001.11653. [Online]. Available: http://arxiv.org/abs/2001.11653 [

5. T. Kojima, T. Nishida, T. Nakamura, A. Tamaoki, A. Hasegawa, Y. Takagi, H. Sato, and K. Ichikawa, ''Keratoconus screening using values derived from auto-keratometer measurements: A multicenter study,'' Amer. J. Ophthalmol., vol. 215, pp. 127–134, Jul. 2020, doi: 10.1016/j.ajo.2020.02.017. [

6. M. D. Twa, S. Parthasarathy, C. Roberts, A. M. Mahmoud, T. W. Raasch, and M. A. Bullimore, ''Automated decision tree classification of corneal shape,'' Optometry Vis. Sci., Off. Publication Amer. Acad. Optometry, vol. 82, no. 12, pp. 1038–1046, Dec. 2005. [

7. C. Fariselli, A. Vega-Estrada, F. Arnalich-Montiel, and J. L. Alio, ''Artificial neural network to guide intracorneal ring segments implantation for keratoconus treatment: A pilot study,'' Eye Vis., vol. 7, no. 1, pp. 1–12, Dec. 2020

8 .S. R. Lin, J. G. Ladas, G. G. Bahadur, S. Al-Hashim, and R. Pineda, ''A review of machine learning techniques for keratoconus detection and refractive surgery screening,'' Seminars Ophthalmol., vol. 34, no. 4, pp. 317–326, 2019, doi: 10.1080/08820538.2019.1620812.

9 .G. W. Armstrong and A. C. Lorch, ''A(eye): A review of current applications of artificial intelligence and machine learning in ophthalmology,'' Int. Ophthalmol. Clinics, vol. 60, no. 1, pp. 57–71, 2020.

10. Chan E, Chong EW, Lingham G, Stevenson LJ, Sanfilippo PG, Hewitt AW, et al. Prevalence of keratoconus based on Scheimpflug imaging: the Raine study. Ophthalmology 2021 Apr;128(4):515-521. [doi: 10.1016/j.ophtha.2020.08.020] [Medline: 32860813]

11.Papali i-Curtin AT, Cox R, Ma T, Woods L, Covello A, Hall RC. Keratoconus prevalence among high school students in New Zealand. Cornea 2019 Nov;38(11):1382-1389. [doi: 10.1097/ICO.0000000000002054] [Medline: 31335534]

12.Ferdi AC, Nguyen V, Gore DM, Allan BD, Rozema JJ, Watson SL. Keratoconus natural progression: a systematic review and meta-analysis of 11 529 eyes. Ophthalmology 2019 Jul;126(7):935-945. [doi: 10.1016/j.ophtha.2019.02.029] [Medline: 30858022]

13.Tuft SJ, Moodaley LC, Gregory WM, Davison CR, Buckley RJ. Prognostic factors for the progression of keratoconus. Ophthalmology 1994 Mar;101(3):439-447. [doi: 10.1016/s0161-6420(94)31313-3] [Medline: 8127564]

14.Pearson AR, Soneji B, Sarvananthan N, Sandford-Smith JH. Does ethnic origin influence the incidence or severity of keratoconus? Eye (Lond) 2000 Aug;14 (Pt 4):625-628. [doi: 10.1038/eye.2000.154] [Medline: 11040911]

15. Wittig-Silva C, Chan E, Islam FM, Wu T, Whiting M, Snibson GR. A randomized, controlled trial of corneal collagen cross-linking in progressive keratoconus: three-year results. Ophthalmology 2014 Apr;121(4):812-821. [doi: 10.1016/j.ophtha.2013.10.028] [Medline: 24393351]

16. Caporossi A, Mazzotta C, Baiocchi S, Caporossi T. Long-term results of riboflavin ultraviolet a corneal collagen cross-linking for keratoconus in Italy: the Siena eye cross study. Am J Ophthalmol 2010 Apr;149(4):585-593. [doi: 10.1016/j.ajo.2009.10.021] [Medline: 20138607]

17. O'Brart DP, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the efficacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross-linkage to halt the progression of keratoconus. Br J Ophthalmol 2011 Nov;95(11):1519-1524. [doi: 10.1136/bjo.2010.196493] [Medline: 21349938]

18. Gore DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. Eye (Lond) 2013 Mar;27(3):329-339 [FREE Full text] [doi: 10.1038/eye.2012.257] [Medline: 23258309]

19. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. J Cataract Refract Surg 2009 Aug;35(8):1358-1362. [doi: 10.1016/j.jcrs.2009.03.035] [Medline: 19631120]

20. Gore DM, Leucci MT, Koay S, Kopsachilis N, Nicolae MN, Malandrakis MI, et al. Accelerated pulsed high-fluence corneal cross-linking for progressive keratoconus. Am J Ophthalmol 2021 Jan;221:9-16. [doi: 10.1016/j.ajo.2020.08.021] [Medline: 32818448]