American Journal of Ophthalmology

Personalized model to predict keratoconus progression from demographic, topographic and genetic data --Manuscript Draft--

Manuscript Number:	AJO-22-105R1
Article Type:	Original Article
Keywords:	Keratoconus; corneal cross-linking; keratoconus genetics; keratoconus prediction
Corresponding Author:	Howard P Maile UCL: University College London UNITED KINGDOM
First Author:	Howard P Maile
Order of Authors:	Howard P Maile
	Ji-Peng Olivia Li
	Mary D Fortune
	Patrick Royston
	Marcello T Leucci
	Ismail Moghul
	Anita Szabo
	Konstantinos Balaskas
	Bruce D Allan
	Alison J Hardcastle
	Pirro Hysi
	Nikolas Pontikos
	Stephen J Tuft
	Daniel M Gore
Abstract:	Purpose : To generate a prognostic model to predict keratoconus progression to corneal cross-linking (CXL). Design: Retrospective cohort study. Methods : We recruited 5025 patients (9341 eyes) with early keratoconus between January 2011 and November 2020. Genetic data from 926 patients was available. We investigated both keratometry or CXL as end-points for progression and used the Royston-Parmar method on the proportional hazards scale to generate a prognostic model. We calculated hazard ratios (HR) for each significant covariate, with explained variation and discrimination, and performed internal-external cross validation by geographic regions. Results: After exclusions, model-fitting comprised 8701 eyes, of which 3232 underwent CXL. For early keratoconus, CXL provided a more robust prognostic model than keratometric progression. The final model explained 33% of the variation in time-to-event: age HR [95% confidence limits] 0.9 [0.90-0.91], maximum anterior keratoconus (n=28) did not significantly contribute to the model. The predicted time-to-event curves closely followed the observed curves during internal-external validation. Differences in discrimination between geographic regions was low, suggesting the model maintained its predictive ability. Conclusions: A prognostic model to predict keratoconus progression could aid patient empowerment, triage and service provision. Age at presentation is the most significant predictor of progression risk. Candidate SNPs associated with keratoconus do not contribute to progression risk.

Suggested Reviewers:	Marcus Ang Marcus.Ang@singhealth.com.sg
	Alfonso lovieno alfonsoiovieno@hotmail.com
	Mani Bhogal manibhogal@aol.co
	Mohammed Ziaei mahdi207@yahoo.com
Opposed Reviewers:	
Response to Reviewers:	Dear Dr Parrish,
	Re: AJO-22-105, "Personalized model to predict keratoconus progression from demographic, topographic and genetic data
	We thank the reviewers for their comments and have responded to each point below. Thank you for considering our revisions.
	I confirm all the authors have reviewed the updated manuscript and responses to reviewers and agree with the changes and responses.
	I had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis as well as the decision to submit for publication. Thank you for your time and for considering our study and we look forward to hearing from you.
	Yours sincerely,
	Dr Daniel Gore Consultant Ophthalmologist in Cornea and External Disease
	Responses to reviewers are found in blue.
	Reviewer #1: The study was well thought out and designed, with appropriate statistica modelling and methodology.
	1. Was multicollinearity considered or assessed in the survival models, as some of th independent variables seem highly collinear?
	Response: Multicollinearity was considered. Correlations between keratometry variables are plotted in Supplemental Figure 2. In particular, K1 and K2 were highly correlated (R2 = 0.91), thus K2 was removed from the model. This is discussed on line 241 with the following text:
	'We chose a model without K2 on the basis of parsimony, which was supported by the fact that K1 and K2 were highly correlated (R2=0.91)'
	2. Are there any differences among prognostic factors across regions?
	Response: We did not identify any significant differences in prognostic factors across regions and we have added this to the text (line 280)
	'We did not identify any significant differences in prognostic factors across regions.'
	3. Table 2: Are there any difference between the univariate and multivariable cohorts

Response:

The cohort used for univariable and multivariable analysis were identical. However, for multivariable analysis, due to the presence of missing data, the number of eyes where all covariates were available was lower than for univariable analysis. We have added this to the manuscript from line 365.

'though the cohort used for univariable and multivariable analysis were identical the number of eyes where all covariates were available was lower than for univariable analysis due to missing data.'

4. How was the prognostic index defined to categorize subjects into 4 risk groups? Also, the text mentioned three risk groups instead of the four in the figures.

Response:

The prognostic index is simply the linear predictor obtained from the Royston Parmar model. For each eye in the cohort, it is a weighted sum of the covariates where the weights are the model regression coefficients. These values were then ordered and divided into quartiles which define 4 risk groups. By splitting into 4 risk groups, we are able to demonstrate group-specific prognosis which helps elucidate the range of discrimination for the model.

We used 4 risk groups for showing model fit and predictive ability (Figures 1 and 3). Figure 2 uses 3 hypothetical patients (not risk groups) to demonstrate the application of the model in a clinical setting.

Reviewer #2: The authors present a very interesting and topical work, which aims to answer the question whether it is possible to predict the progression of keratoconus in the examined eyes on the basis of the measured parameters, which subsequently leads to the necessity of CXL procedure.

The paper analyses the results of the measured parameters on a very high number of eyes (8701), which enhances the relevance of the statistical evaluation of these results, and also analyses the risk factors from the genomic point of view of some of the examined patients.

5. The study points to variable factors such as age, keratometry, gender (evident from the results in tables) as decisive factors at the time of the first examination of the patient as determinants for the "probable" necessity of the procedure. I would recommend address the concrete values of these parameters in the text also - it would be more clearer and more didactic for the reader of the paper if the authors also commented and analysed these results in the text (which are obvious from their detailed tables) and determined at least indicatively, e.g. specific limits of the patient's age and specific keratometry values, according to which practical ophthalmologists could inform the patient at the initial visit about the possible further course of the disease and the likelihood of needing CXL procedure.

Response:

We have added from line 311

"Of the significant covariates in our model, younger age made the greatest contribution to our model. Thus, one should have a lower threshold for treatment in younger patients."

We have not provided specific limits on the patient's age and keratometry as it is beyond the scope of this study. We have referenced the treatment criteria that we have applied. (Gore et al 2021) and from line 355 we have added:

'The effect of recommendations based on this model on clinical practice is yet to be evaluated.'

Reviewer #3: Dear authors,

Establishing more robust criteria for patient selection and prompt CXL treatment is one of the most challenging issues in keratoconus management. This manuscript aims at identifying and classifying the variables influencing keratoconus progression. The authors identified age, Kmax and minimum corneal thickness as the most significant

covariates in establishing keratoconus progression, genetic or geographic covariates did not improve the accuracy of their model. Based upon the results, the conclusions of the work are not new. However, the manuscript may deserve further revision in view of the extensive work done by the authors.

6. The weakest point of the manuscript is common to all other publications in the topic, which is the definition of keratoconus progression. The authors should discuss this issue further by revising the most recent literature in the topic.

Response:

In this study, we have chosen to use crosslinking as a surrogate for progression. This is a meaningful endpoint for patients and does not rely simply on meeting defined keratometric thresholds. We recognise that there is variability in the definition of keratoconus progression in the literature and have included some updated references on this point, which also explores the increased variability in more advanced disease, including Hashemi et al 2022, Ozalp et al 2022, Flockerzi et al 2022 and Kreps et al 2020.

We have clarified these points in the text from line 63.

'We considered the date of numeric progression, as well as the date when CXL was performed, as alternative end-points to define keratoconus progression. Although the use of keratometry as an end-point may appear the more objective method, there is variability on the definition of progression reported in the literature and conclusions may vary with the definition that is adopted.(Vinciguerra et al. 2021; Shajari et al. 2019; Ozalp and Atalay 2022; Hashemi et al. 2022) Repeatability thresholds are not usually tailored to individual eyes (i.e. an increase in Kmax by 1 D is not significant in advanced keratoconus) although there is growing evidence on the variability of measurements in more advanced disease and the need for tailoring numerical progression definitions to the disease state, and distinguishing real progression from inherent measurement variability.(Flynn et al. 2016; Flockerzi et al. 2022; Kreps et al. 2020) '

7. I recommend to delete from analysis both the genetic and geographic data. Genetic data were collected from a small subgroup of patients and did not show to hold significant information. The authors may prefer to analyze genetic data further in a separate work. Geographic data remains unclear; they were not related to ethnicity or relative marriage or migration issues. The authors may prefer to delete these information and Figure 2 (erroneously described as figure legend 3) from the manuscript.

Response:

Thank you for this suggestion. However, we believe that there is value in the genetic report. Assessing the impact of SNPs associated with keratoconus on risk of progression is novel and builds on the recent findings of genetic loci significantly associated with keratoconus. (Hardcastle et al 2021 Commun Biol.) Whilst genetic data was only available for a subgroup, it is still a sizeable subgroup out of a very large number of eyes. Also, our analysis puts an upper bound on the contribution of SNPs associated with keratoconus on its progression.

The geographic data was split to perform internal external validation and is a recognised method of validating the model with the original dataset. Therefore, they are not intended to relate to ethnicity or other variables.

We have now correctly labelled Figure 2.

8. The authors could not recommend any clinical practice based upon a mathematical model. The model shall be validated before designing any preferred best practice in keratoconus management. For example, validation shall be performed by independent clinical centers investigating a cohort of keratoconus patients for at least 2 years to assess the performance of the model.

Response: Thank you for this suggestion. We performed internal validation but have not tested the

generalisability of the model with an external dataset. Therefore, we have not proposed any practice guidance. We plan to test the model before making recommendations on clinical practice based on the model predictions. We have clarified this recommendation in the Discussion (line 355).

'Recommendations based on this model on clinical practice is yet to be evaluated.'

9. The authors shall discuss their results in comparison with the state-of-the-art knowledge, showing common points and, if any, main differences.

Response:

We have included comparison with other studies, including Quartilho et al from line 326, and notably Kato et al 2020 and Kato et al 2021 comparing the baseline factors as predictors of crosslinking from line 332, where a deep learning algorithm was utilised.

'Using logistic regression, Kato et al. found that the two strongest factors associated with the requirement for CXL were age and Kmax, which is consistent with our findings. Moreover, their team also found that with deep learning age combined with corneal tomography maps could predict progression and the need for crosslinking. (Kato et al. 2021)'

10. Figure 3 (erroneously described as figure legend 2) is unclear. Please add the equation and the 95% CI to each curve.

Response:

We have added 95% CI to each curve as suggested and added details of the equation to the figure text as follows from line 520:

'The equation used to generate the curves is: $S(t)=^{-(-H(t))}$, where H(t) is the cumulative hazard function and is commonly expressed as $ln(H(t))=s(ln(t))+x\beta$, where s(ln(t)) is a restricted cubic spline function of log time, β is the vector of coefficients and x is the vector of covariates.'

We have also corrected the labelling of the figures.

Other comments:

11. Please use the standardized term "corneal cross-linking" and not "collagen corneal cross-linking".

Response:

This has been corrected throughout the text.

12. The 98% success rate is too optimistic for current state-of-the-art CXL. As the authors may find in meta-analysis studies, the range of CXL treatment efficacy spans between 10% and 90%; there is high variability caused by the plethora of riboflavin dosing and UV-A irradiation protocols.

Response:

We agree there is variability in reported outcomes related to treatment protocols and also to the definition of failure. The 98% success rate is our own data from our 2-year outcomes report (Ref 7, Gore et al 2021). We have added a range for other publications (88% - 100%) from line 45.

⁶Corneal cross-linking (CXL) by topical application of riboflavin, followed by irradiation with UV-A light, can arrest progression of keratoconus in up to 88% to 100% of eyes even when there is relatively advanced disease.2–6

±

Abstract

Purpose: To generate a prognostic model to predict keratoconus progression to corneal cross-linking (CXL).

Design: Retrospective cohort study.

Methods: We recruited 5025 patients (9341 eyes) with early keratoconus between January 2011 and November 2020. Genetic data from 926 patients was available. We investigated both keratometry or CXL as end-points for progression and used the Royston-Parmar method on the proportional hazards scale to generate a prognostic model. We calculated hazard ratios (HR) for each significant covariate, with explained variation and discrimination, and performed internal-external cross validation by geographic regions.

Results: After exclusions, model-fitting comprised 8701 eyes, of which 3232 underwent CXL. For early keratoconus, CXL provided a more robust prognostic model than keratometric progression. The final model explained 33% of the variation in time-to-event: age HR [95% confidence limits] 0.9 [0.90-0.91], maximum anterior keratometry (Kmax) 1.08 [1.07-1.09], and minimum corneal thickness 0.95 [0.93-0.96] as significant covariates. Single nucleotide polymorphisms (SNPs) associated with keratoconus (n=28) did not significantly contribute to the model. The predicted time-to-event curves closely followed the observed curves during internal-external validation. Differences in discrimination between geographic regions was low, suggesting the model maintained its predictive ability.

Conclusions: A prognostic model to predict keratoconus progression could aid patient empowerment, triage and service provision. Age at presentation is the most significant predictor of progression risk. Candidate SNPs associated with keratoconus do not contribute to progression risk.





Moorfields Eye Hospital NHS Foundation Trust

162 City Road, London EC1V 2PD Phone: 020 7253 3411

March 31st, 2022

Editor in Chief American Journal of Ophthalmology

Dear Dr Parrish,

Re: AJO-22-105, "Personalized model to predict keratoconus progression from demographic, topographic and genetic data

We thank the reviewers for their comments and have responded to each point below. Thank you for considering our revisions.

I confirm all the authors have reviewed the updated manuscript and responses to reviewers and agree with the changes and responses.

I had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis as well as the decision to submit for publication. Thank you for your time and for considering our study and we look forward to hearing from you.

Yours sincerely,

Dr Daniel Gore Consultant Ophthalmologist in Cornea and External Disease

Patron: Her Majesty The Queen Chairman: Tessa Green, CBE Chief executive: David Probert www.moorfields.nhs.uk

Responses to reviewers are found in blue.

Reviewer #1: The study was well thought out and designed, with appropriate statistical modelling and methodology.

1. Was multicollinearity considered or assessed in the survival models, as some of the independent variables seem highly collinear?

Response:

Multicollinearity was considered. Correlations between keratometry variables are plotted in Supplemental Figure 2. In particular, K1 and K2 were highly correlated ($R^2 = 0.91$), thus K2 was removed from the model. This is discussed on line 241 with the following text:

'We chose a model without K2 on the basis of parsimony, which was supported by the fact that K1 and K2 were highly correlated (R²=0.91)'

2. Are there any differences among prognostic factors across regions?

Response:

We did not identify any significant differences in prognostic factors across regions and we have added this to the text (line 280)

'We did not identify any significant differences in prognostic factors across regions.'

3. Table 2: Are there any difference between the univariate and multivariable cohorts?

Response:

The cohort used for univariable and multivariable analysis were identical. However, for multivariable analysis, due to the presence of missing data, the number of eyes where all covariates were available was lower than for univariable analysis. We have added this to the manuscript from line 365.

'though the cohort used for univariable and multivariable analysis were identical the number of eyes where all covariates were available was lower than for univariable analysis due to missing data.'

4. How was the prognostic index defined to categorize subjects into 4 risk groups? Also, the text mentioned three risk groups instead of the four in the figures.

Response:

The prognostic index is simply the linear predictor obtained from the Royston Parmar model. For each eye in the cohort, it is a weighted sum of the covariates where the weights are the model regression coefficients. These values were then ordered and divided into quartiles which define 4 risk groups. By splitting into 4 risk groups, we are able to demonstrate group-specific prognosis which helps elucidate the range of discrimination for the model.

We used 4 risk groups for showing model fit and predictive ability (Figures 1 and 3). Figure 2 uses 3 hypothetical patients (not risk groups) to demonstrate the application of the model in a clinical setting.

Reviewer #2: The authors present a very interesting and topical work, which aims to answer the

question whether it is possible to predict the progression of keratoconus in the examined eyes on the basis of the measured parameters, which subsequently leads to the necessity of CXL procedure. The paper analyses the results of the measured parameters on a very high number of eyes (8701), which enhances the relevance of the statistical evaluation of these results, and also analyses the risk factors from the genomic point of view of some of the examined patients.

5. The study points to variable factors such as age, keratometry, gender (evident from the results in tables) as decisive factors at the time of the first examination of the patient as determinants for the "probable" necessity of the procedure. I would recommend address the concrete values of these parameters in the text also - it would be more clearer and more didactic for the reader of the paper if the authors also commented and analysed these results in the text (which are obvious from their detailed tables) and determined at least indicatively, e.g. specific limits of the patient's age and specific keratometry values, according to which practical ophthalmologists could inform the patient at the initial visit about the possible further course of the disease and the likelihood of needing CXL procedure.

Response:

We have added from line 311

"Of the significant covariates in our model, younger age made the greatest contribution to our model. Thus, one should have a lower threshold for treatment in younger patients."

We have not provided specific limits on the patient's age and keratometry as it is beyond the scope of this study. We have referenced the treatment criteria that we have applied. (Gore et al 2021) and from line 355 we have added:

'The effect of recommendations based on this model on clinical practice is yet to be evaluated.'

Reviewer #3: Dear authors,

Establishing more robust criteria for patient selection and prompt CXL treatment is one of the most challenging issues in keratoconus management. This manuscript aims at identifying and classifying the variables influencing keratoconus progression. The authors identified age, Kmax and minimum corneal thickness as the most significant covariates in establishing keratoconus progression, genetic or geographic covariates did not improve the accuracy of their model. Based upon the results, the conclusions of the work are not new. However, the manuscript may deserve further revision in view of the extensive work done by the authors.

6. The weakest point of the manuscript is common to all other publications in the topic, which is the definition of keratoconus progression. The authors should discuss this issue further by revising the most recent literature in the topic.

Response:

In this study, we have chosen to use crosslinking as a surrogate for progression. This is a meaningful endpoint for patients and does not rely simply on meeting defined keratometric thresholds. We recognise that there is variability in the definition of keratoconus progression in the literature and have included some updated references on this point, which also explores the increased variability in more advanced disease, including Hashemi et al 2022, Ozalp et al 2022, Flockerzi et al 2022 and Kreps et al 2020.

We have clarified these points in the text from line 63.

'We considered the date of numeric progression, as well as the date when CXL was performed, as alternative end-points to define keratoconus progression. Although the use of keratometry as an end-point may appear the more objective method, there is variability on the definition of progression reported in the literature and conclusions may vary with the definition that is adopted. (Vinciguerra et al. 2021; Shajari et al. 2019; Ozalp and Atalay 2022; Hashemi et al. 2022) Repeatability thresholds are not usually tailored to individual eyes (i.e. an increase in Kmax by 1 D is not significant in advanced keratoconus) although there is growing evidence on the variability of measurements in more advanced disease and the need for tailoring numerical progression definitions to the disease state, and distinguishing real progression from inherent measurement variability.(Flynn et al. 2016; Flockerzi et al. 2022; Kreps et al. 2020) '

7. I recommend to delete from analysis both the genetic and geographic data. Genetic data were collected from a small subgroup of patients and did not show to hold significant information. The authors may prefer to analyze genetic data further in a separate work. Geographic data remains unclear; they were not related to ethnicity or relative marriage or migration issues. The authors may prefer to delete these information and Figure 2 (erroneously described as figure legend 3) from the manuscript.

Response:

Thank you for this suggestion. However, we believe that there is value in the genetic report. Assessing the impact of SNPs associated with keratoconus on risk of progression is novel and builds on the recent findings of genetic loci significantly associated with keratoconus. (Hardcastle et al 2021 Commun Biol.) Whilst genetic data was only available for a subgroup, it is still a sizeable subgroup out of a very large number of eyes. Also, our analysis puts an upper bound on the contribution of SNPs associated with keratoconus on its progression.

The geographic data was split to perform internal external validation and is a recognised method of validating the model with the original dataset. Therefore, they are not intended to relate to ethnicity or other variables.

We have now correctly labelled Figure 2.

8. The authors could not recommend any clinical practice based upon a mathematical model. The model shall be validated before designing any preferred best practice in keratoconus management. For example, validation shall be performed by independent clinical centers investigating a cohort of keratoconus patients for at least 2 years to assess the performance of the model.

Response:

Thank you for this suggestion. We performed internal validation but have not tested the generalisability of the model with an external dataset. Therefore, we have not proposed any practice guidance. We plan to test the model before making recommendations on clinical practice based on the model predictions. We have clarified this recommendation in the Discussion (line 355).

'Recommendations based on this model on clinical practice is yet to be evaluated.'

9. The authors shall discuss their results in comparison with the state-of-the-art knowledge, showing common points and, if any, main differences.

Response:

We have included comparison with other studies, including Quartilho et al from line 326, and notably Kato et al 2020 and Kato et al 2021 comparing the baseline factors as predictors of crosslinking from line 332, where a deep learning algorithm was utilised.

'Using logistic regression, Kato et al. found that the two strongest factors associated with the requirement for CXL were age and Kmax, which is consistent with our findings. Moreover, their team also found that with deep learning age combined with corneal tomography maps could predict progression and the need for crosslinking. (Kato et al. 2021)'

10. Figure 3 (erroneously described as figure legend 2) is unclear. Please add the equation and the 95% CI to each curve.

Response:

We have added 95% CI to each curve as suggested and added details of the equation to the figure text as follows from line 520:

'The equation used to generate the curves is: $S(t) = e^{-H(t)}$, where H(t) is the cumulative hazard function and is commonly expressed as $ln(H(t)) = s(ln(t)) + x\beta$, where s(ln(t)) is a restricted cubic spline function of log time, β is the vector of coefficients and x is the vector of covariates.'

We have also corrected the labelling of the figures.

Other comments:

11. Please use the standardized term "corneal cross-linking" and not "collagen corneal cross-linking".

Response:

This has been corrected throughout the text.

12. The 98% success rate is too optimistic for current state-of-the-art CXL. As the authors may find in meta-analysis studies, the range of CXL treatment efficacy spans between 10% and 90%; there is high variability caused by the plethora of riboflavin dosing and UV-A irradiation protocols.

Response:

We agree there is variability in reported outcomes related to treatment protocols and also to the definition of failure. The 98% success rate is our own data from our 2-year outcomes report (Ref 7, Gore et al 2021). We have added a range for other publications (88% - 100%) from line 45.

'Corneal cross-linking (CXL) by topical application of riboflavin, followed by irradiation with UV-A light, can arrest progression of keratoconus in up to 88% to 100% of eyes even when there is relatively advanced disease. $\frac{2-6}{2}$

<u>±</u>

	1	
1 2	2	
3 4	3	
5 6	4	
7	5	Personalized model to predict keratoconus progression from
8 9	6	demographic, topographic and genetic data
10 11	7	
12 13	8	
14 15	9	Short title: Keratoconus progression
16 17	10	
18	11	
19 20	12	
21 22	13	Howard P Maile¹ ⁻ ; Ji-Peng Olivia Li² ⁻ ; Mary D Fortune³; Patrick Royston⁴;
23 24	14	Marcello T Leucci²; Ismail Moghul¹; Anita Szabo¹; Konstantinos Balaskas²;
25 26	15	Bruce D Allan ² ; Alison J Hardcastle ¹ ; Pirro Hysi ^{5,6} ; Nikolas Pontikos ¹ *; Stephen J
27 28	16	Tuft²*; Daniel M Gore²*\$
29	17	
30 31	18	
32 33	19	¹ UCL Institute of Ophthalmology, 11-43 Bath Street, London EC1V 9EL
34 35	20	² Moorfields Eye Hospital NHS Foundation Trust, 162 City Road, London EC1V
36 37	21	2PD
38 39	22	³ MRC Biostatistics Unit, Cambridge Institute of Public Health, University of
40	23	Cambridge, UK
41 42	24	^₄ MRC Clinical Trials Unit at UCL, Aviation House, 125 Kingsway, London,
43 44	25	WC2B 6NH
45 46	26	Section of Ophthalmology, School of Life Course Sciences, King's College
47 48	27	London
49 50	28	⁶ Department of Twin Research and Genetic Epidemiology, King's College
51	29	London
52 53	30	
54 55	31	* Authors contributed equally
56 57	32	^s Corresponding author: Mr Daniel Gore, Moorfields Eye Hospital, 162 City
58 59	33	Road, London, EC1V 2PD, UK, <u>daniel.gore1@nhs.net</u> , 0207 253 3411.
60 61	34	
62		
63		

35 Keywords: Keratoconus, corneal cross-linking, keratoconus genetics,

36 keratoconus prediction

Page Break

39 Introduction

Keratoconus is a common corneal ectasia that causes irregular astigmatism, scarring and loss of vision. Thinning and steepening can progress through childhood and early adulthood, but the shape of most eyes stabilizes by the third or fourth decade. Without intervention, keratoconus can lead to severe visual loss, with approximately 10% of eyes eventually requiring corneal transplantation.¹ Corneal cross-linking (CXL) by topical application of riboflavin, followed by irradiation with UV-A light, can arrest progression of keratoconus in up to 88% to 100% of eyes even when there is relatively advanced disease.²⁻⁶ The potential benefit of CXL is to prevent visual deterioration with a relatively low risk procedure that is cost effective for healthcare providers.⁷⁻⁹ However, CXL is usually not offered to all patients at presentation because the disease may have already stabilized. In the recent KERALINK study 43% of children <17 years of age at presentation had not progressed after 18 months.¹⁰ The definition of progression also varies with the severity of keratoconus, but for early disease a common threshold is either an increase in the maximum keratometry (Kmax) of >1 dioptre, a change in the manifest refractive spherical equivalent of >0.50 dioptre, or an increase in manifest refractive cylinder of >1 dioptre.^{2,11} Depending on the rate of progression this threshold may be passed in a few months, years, or not at all. At the first assessment it can be a challenge to distinguish eyes that are at risk of rapid progression from those where it is safe to monitor. Unnecessary review visits are a burden to the patient and the care system.

We considered the date of numeric progression,⁶ as well as the date when CXL
was performed, as alternative end-points to define keratoconus progression.
Although the use of keratometry as an end-point may appear the more
objective method, there is variability on the definition of progression reported in
the literature and conclusions may vary with the definition that is adopted.¹¹⁻¹⁴
Repeatability thresholds are not usually tailored to individual eyes (i.e. an

increase in Kmax by 1 D is not significant in all eyes) although there is growing evidence on the variability of measurements in more advanced disease and the need for tailoring numerical progression definitions to the disease state, and distinguishing real progression from inherent variability of measurement modalities.¹⁵⁻¹⁷ Finally, patients who receive CXL prior to progression must be censored from the dataset even though these eyes are likely to have been at risk of progression. This type of informative censoring creates a bias.¹⁰ In contrast, the time to CXL depends on several variables that include numeric disease progression, but also incorporates patient-specific risk factors for future progression. Its strength is that it is an easily comprehensible and meaningful end-point for patients. It encompasses individual risk factors that are not considered when imaging is used in isolation and it has been used by others as defining the event of interest.¹⁹

For these reasons we have used demographic and serial tomography data from a large cohort of patients to generate a time-to-event model to predict the probability of an individual progressing to CXL. Because the Cox proportional hazards method does not generate smooth time-to-event curves, we used the Royston-Parmar model to achieve direction estimates of the hazard function.²⁰ We also performed a further analysis of a subset of patients who had genetic data in the form of single-nucleotide polymorphisms (SNP) generated as part of a study to determine keratoconus risk.²¹

92 Methods

93 Cohort

The study protocol was reviewed and approved by the Clinical Audit Assessment Committee of Moorfields Eye Hospital NHS Foundation Trust (reference CA17/CED/03). Institutional Review Board (IRB) approval was obtained and individual patient consent was not required. The study conformed to the tenets of the Declaration of Helsinki. We identified from the Moorfields Eye Hospital electronic health record database (OpenEyes) patients aged 13 years and above diagnosed with clinical or suspected keratoconus who attended our Early Keratoconus Clinic (EKC) between January 2011 and November 2020. Clinical data included keratometry (Kmax, Front K1, Front K2,

Back K1, Back K2), and pachymetry (minimum corneal thickness) captured by Scheimpflug tomography (Pentacam HR, Oculus GmbH, Wetzlar). We only included scans with a quality score of 'good' or 'ok', and where multiple scans were taken on the same day we used the mean value. The date of all CXL procedures was recorded. The protocol for offering CXL throughout the study period was, i) a documented history prior to referral to the EKC of our hospital of significant recent disease progression, ii) a change in contemporary measurements of 95% above the repeatability limits of the baseline measurements as shown in Supplemental Table 1 (available at http://www.ajo.com),⁶ or iii) a patient considered by a clinician to be at high risk of progression despite their not fulfilling the above two criteria. Exclusion criteria included pregnancy or breastfeeding, uncontrolled ocular surface disease or a minimum corneal thickness less than 375 µm.

All the data used for model fitting started from the first appointment in the EKC. Patient demographics included age, gender, smoking status (current or ex/non-smoker) ethnicity and postcode. Ethnicity was coded as 1 for 'Black' or 'South Asian or South Asian British' and 0 for any other category (excluding missing values). Before model fitting, the pachymetry in microns was divided by 10 to generate a meaningful scale. For the primary analysis, eyes with any missing data were excluded. We also explored multiple imputation, which avoids data exclusion by generating multiple versions of the dataset with missing values replaced with values sampled from an appropriate distribution. To see whether genetic data can help predict keratoconus progression, we used 28 candidate SNPs from a recent keratoconus genome-wide association study that contained 926 patients from Moorfields Eye Hospital.²¹ The SNP data was encoded as either 0 (homozygous reference genotype), 1 (heterozygous genotype), or 2 (homozygous variant genotype). We chose to use an additive encoding, thus the risk of disease increases additively with the degree of genetic variation.²² Anonymized data were then exported to Excel software for analysis (version 15.24 2016, Microsoft Corp.).

135 Model Fitting and Covariate Selection

A Royston-Parmar flexible parametric survival model was fitted to the data to predict the probability of an eye progressing to CXL.²³ Initial analysis of the covariates was performed by univariate analysis using the same model characteristics as the multivariable model. When selecting covariates for the final multivariable model, we used backwards stepwise selection with a significance level of 0.05. We used linear covariates for ease of interpretation of our final model. To create a more parsimonious model we examined the effect on explained variation and discrimination of removing single variables from the model.

146 Keratometric Progression Sensitivity Analysis

We included a sensitivity analysis in which we investigated keratometric progression as an alternative end-point. Keratometric progression was defined using thresholds from Gore et al 2021.⁶ When using numerical thresholds to define progression, the appointments for eyes beyond the date of CXL cannot be used. However, censoring these eyes at the date of CXL represents informative censoring. Based on the recommendations of Clarke et al¹⁰ for investigating the impact of informative censoring, we generated a 'best case' dataset where eyes were censored at the CXL date and a 'worst case' dataset where patients were assumed to progress at the CXL date. The corresponding Kaplan Meier curves were plotted to provide a visual comparison of the two datasets. A Royston-Parmar model was then fitted on both datasets. We used the same techniques (backwards stepwise selection, significance level of 0.05) as described in the previous section to fit the model and compare the explained variation and hazard ratios.

162 Multivariable Model Validation

We validated the model using internal-external cross validation in which we split the dataset by geographical region.^{24,25} For the kth region, the model is fitted on the full dataset excluding region k and then Kaplan-Meier curves and predicted survival curves were generated for region k. Seven geographical regions were created based on the patient's postcode as shown in Supplemental Figure 1 (available at http://www.ajo.com). To quantitatively assess the validation, Royston and Sauerbrei's D statistic was calculated for both the model fitted

170 from data excluding region k ($D_{(k)}$) and also the model applied to region k (D_k).²⁶ 171 The difference between these two discrimination metrics (D_k - $D_{(k)}$) was calculated 172 with its corresponding standard error to assess the predictive ability of the 173 model. To demonstrate how the model could be used in practice, we include 174 three hypothetical patients' eyes with different progression risk profiles (high, 175 medium, low risk) and plot the predicted time-to-event curve for each shown in 176 Figure 2.

178Statistical Analysis

The event of interest was defined as the date that the eye underwent CXL. We calculated the time-to-event as the difference between the first appointment in our service and the date of CXL (or the last patient appointment in the case of censoring). Since we had paired observations (eyes), we used variancecorrected models to account for correlation between eyes and to ensure that robust standard errors were produced. The choice of scale and selection of degrees of freedom for the Royston-Parmar model was informed by inspecting the Akaike information criterion (AIC) and Bayes information criterion (BIC)²⁰ and the results of this were balanced with ease of interpretation. See Supplemental Table 2 and Supplemental Text 1 (available at http://www.ajo.com) for further explanation. Royston and Sauerbrei's D statistic was used as a measure of discrimination and R²_D as a measure of explained variation (both calculated on the natural scale of the model). Although all of the primary results were generated from a complete case analysis, we performed an additional analysis using multiple chained imputation (predictive mean matching approach with 5 nearest neighbors). Model fitting was performed in Stata 13 (StataCorp LP, Texas, USA) and the Royston-Parmar model was fitted using the stpm2 package from Stata 13.

198 Results

Cohort

From a potential of 9,341 eyes (4316 pairs of eyes and 709 individual eyes), the
final model used 8,701 eyes of 4,823 patients, with 3,232 eyes that had CXL.

- 202 The mean age was 28.3 years with standard deviation of 7.1 years. We
- 203 excluded 640 eyes with missing data. Table 1 summarizes the available

204 covariates along with missing data percentages. See Supplemental Text 2 and
205 Supplemental Table 3 (available at http://www.ajo.com) for a description of the
206 multiple imputation results.

208 Model Fitting and Covariate Selection (Genetic Data)

We analyzed patients with genetic data separately because this data was only available for ~14% of patients. Of 926 patients (1852 eyes) with genetic data, 531 eyes were excluded with incomplete keratometry or CXL data, which left 1321 eyes, of which 665 had CXL. With univariate analysis of the 28 SNPs only rs72631889 was found to be significant (P=0.01) (Supplemental Table 4 (available at http://www.ajo.com)). We then produced a multivariable model via backwards selection on this subset of eyes using corneal data, patient data and rs72631889 as an additional covariate as shown in Supplemental Table 5 (available at http://www.ajo.com). However rs72631889, although significant (P=0.005), had a negligible contribution (0.3%) to the explained variation in the final model.

221 Model Fitting and Covariate Selection (Excluding Genetic Data)

The results of the univariate time-to-event analysis on the hazards scale using a Royston-Parmar flexible parametric model is shown in Table 2. Genetic data was excluded from this analysis. All variables except smoking status were significant. The explained variation (R²_D) and discrimination (D) were highest for age (17%) and Kmax (15%) with Front K1, Front K2, Back K1, Back K2 and pachymetry each explaining 6-10% of the variation. Notably, gender and ethnicity, although significant in the univariate analysis, did not contribute to explained variation. The hazard ratios for significant covariates indicate that increasing age at presentation, greater pachymetry and flatter (less negative) posterior keratometry values decrease risk of having CXL, whilst steeper anterior keratometry values and male gender increase the risk of having CXL. When we fitted a multivariable model the significant covariates were age,

236 Kmax, Front K1, Front K2 and pachymetry (Table 2). When we removed single

237 variables from the model the effect this had on explained variation and

discrimination is shown in Supplemental Table 6 (available at http://www.ajo.com). Age was the most important covariate (16.7%), with Kmax contributing ~5% of explained variation. K1, K2 and pachymetry had a small effect (<1%) when removed individually. We chose a model without K2 on the basis of parsimony, which was supported by the fact that K1 and K2 were highly correlated (R²=0.91) as shown in Supplemental Figure 2 (available at http://www.ajo.com). The final fitted model hazard ratios can be seen on the multivariable column of Table 2. It is notable that an increase in K1 now has a protective effect in the final model. The explained variation and discrimination for the final model were 32.7% and 1.43 respectively.²⁷ The opposing effect of Kmax and Front K1 can be explained by examining their regression coefficients before converting to hazard ratios; Kmax has a positive coefficient (0.0795) and Front K1 has a negative coefficient (-0.0749). This is logically similar to including the combined covariate (Kmax - Front K1) in the model which can be viewed clinically as a proxy for irregular astigmatism. We also investigated combining K1 and K2 into a single covariate as K2-K1 (standard definition of astigmatism), but the corresponding p value was not significant.

Figure 1 visually depicts the result of applying the final model to the original
dataset. As expected, the predicted mean survival curves closely follow the
Kaplan-Meier curves. To demonstrate the use of the model in clinical practice,
survival curves for three hypothetical patients followed for five years are shown
in Figure 2. We have also produced a web application from the model which
can be accessed at http://beta.moorfieldscxl.com.

263 Keratometric Progression Sensitivity Analysis

The results of the keratometric progression sensitivity analysis can be found in the Supplementary Material. By examining the Kaplan Meier curves in Supplemental Figure 3, we can see that the best case time-to-event curve indicates a 40% survival probability at 5 years whilst the worst case curve indicates a 27% survival probability at 5 years. This 13% difference in survival probability at 5 years represents the upper bound of the discrepancy in survival probability within the data. After fitting the Royston-Parmar model, amongst the hazard ratios which overlap (age, Kmax, k2), there was reasonable similarity

(Supplemental Tables 8 and 9). Most importantly, the model fitted to the best
case had an explained variation of 11% compared to 23% for the worst case
indicating a significant difference in model performance depending on the
assumptions used for handling eyes which received CXL.

277 Multivariable Model Validation

When performing validation using internal-external cross validation, Figure 3 shows the ability of our final model to predict keratoconus progression across different geographic regions. We did not identify any significant differences in prognostic factors across regions. The model prediction curves generally follow the Kaplan Maier curves. Notably, region 5 (South West Greater London) and region 7 (other regions) have a worse predictive performance than the other regions, indicating that these regions have different characteristics compared with the remainder of the dataset used for model fitting. This could be due to differing patient characteristics, such as complex cases that required referral to our tertiary referral centre rather than being managed locally. Overall, the prediction becomes less accurate over time, which is expected due to low numbers with follow-up beyond three years. Supplemental Table 7 displays quantitative validation results of the model using internal external validation. The difference column $D_k - D_{(k)}$ is a measure of predictive ability. Region 7 (other regions outside of Greater London) has the greatest discrepancy in discrimination (-0.26) which indicates that the model fitted when excluding region 7 had greater discriminative ability than when applied to region 7 alone.

Discussion

In this study we have incorporated demographic, keratometric, and genetic data to generate a prognostic model of keratoconus progression to CXL. We have shown that parameters recorded at the first examination (age, Kmax, Front K1, minimum pachymetry) can produce a time-to-event curve to calculate a personalized risk for keratoconus progression. Although we chose time to CXL rather than keratometric progression as the end point for the time-to-event analysis, we performed a sensitivity analysis using keratometric progression, and found that a CXL model accounts for a much higher proportion of the

explained variation (33%) compared to the keratometric model (11% or 23% for best and worst case respectively). The opposing effects of Kmax and Front K1 were unexpected, but similar to including the combined covariate (Kmax - Front K1) in the model; a possible explanation is that the opposing effect is the result of an increase in irregular astigmatism. Of the significant covariates in our model, younger age made the greatest contribution to our model. Thus, one should have a lower threshold for treatment in younger patients. When applying internal-external cross validation, the survival curves closely followed the Kaplan Meier survival curves for each of the geographic regions, which indicates generalisability, and model discrimination between training and cross validation groups was similar, indicating that the predictive ability is well

maintained. Finally, our SNP genetic data had limited additional predictive utility
for keratoconus progression. However, the genetic dataset was relatively small
(926 patients), and recruitment was based on the presence of keratoconus, as
opposed to the severity of keratoconus, or any other index of risk of rapid
progression.

The Royston-Parmar model has previously been used to predict the likelihood of the worst eye of patients with keratoconus progressing to corneal transplantation.²⁸ In their final model, Quartilho et al chose 3 significant covariates: Kmax, age and ethnicity. The reported covariate hazard ratios that overlap with our study (Kmax and age) were different in magnitude but in the same direction. When performing internal validation their model exhibited good predictive ability. They produced time-dependent receiver operating curves using the validation set and found one-year sensitivity and specificity to be 92.8% and 94.6% respectively. Using logistic regression, Kato et al. found that the two strongest factors associated with the requirement for CXL were age and Kmax, which is consistent with our findings.¹⁹ Moreover the team went on to find that age combined with corneal tomography maps was able to predict progression and need for crosslinking using deep learning.²⁹

An ability to generate personalized time-to-event curves that predict
progression to CXL (Figure 2) could directly inform clinical decisions that benefit
patient care. Firstly, patients may better understand their own risk for

progression and feel more confident in choosing their management options. Secondly, for both clinicians and patients, the prediction of progression may contribute to scheduling treatments, including prioritizing patients at high risk of early progression. For example, patients at high risk with a 98% probability of progressing to CXL at 5 years could be offered CXL at the point of first diagnosis without waiting to demonstrate keratometric progression. Medium risk patients may benefit from a period of clinician-led topographic monitoring. For the lowest risk patients, optometry-led monitoring in the community may be sufficient. This risk stratification could be tailored to regions and reflect local needs and resources such as provision of monitoring services in regions with lower risk and greater capacity for CXL in areas with more high risk patients. Finally, when a decision is made to postpone CXL for further monitoring, the time-to-event curve can contribute to decisions on the scheduling of future follow up reviews, with perhaps shorter time periods where the curve is steepest. Recommendations based on this model on clinical practice is yet to be evaluated.

Our study is subject to several limitations inherent to our dataset. First, if patients had CXL at another hospital, this may not be reliably recorded in the source database. This could lead to a very small number of patients being included in the analysis who have already had CXL. Second, ethnicity is a well established risk factor for keratoconus and keratoconus progression, 27,30,31 but ethnicity is now an optional field at patient registration at our institution and this information was unavailable for approximately 50% of our dataset. However, even when we restricted the dataset to those with ethnicity records, it was not found to be a significant covariate. Third, though the cohort used for univariable and multivariable analysis were identical the number of eyes where all covariates were available was lower than for univariable analysis due to missing data. Finally, when we used multiple imputation to generate a multivariable model, ethnicity was still not found to be significant. In the model fitting process we chose to use a simple backwards selection as opposed to the multivariate fractional polynomial (MFP) method.³² In our initial investigations, the results of MFP yielded nonlinear functional forms of the covariates and, whilst this method may have slightly increased the predictive power of the

374 prognostic model, the resulting hazard ratios would be very hard to interpret. In 375 addition, we did not examine time dependent effects for the covariates, which 376 may provide a more accurate model fit, and future studies should examine this 377 option. Finally, although no external validation dataset was available, internal 378 external cross validation allowed us to confirm that our model is generalizable 379 across geographical regions.

 In conclusion, we have fitted a prognostic model for progression of keratoconus to CXL which generates a time-to-event curve using age, Kmax, Front K1, minimum pachymetry from time of presentation. Incorporation of a relatively small genetic dataset does not improve the explained variation of our model. Personalized modeling of risk may improve patients' understanding of their condition and the need for CXL. Such a model may help better improve patients and aid clinician decision making to CXL to achieve better outcomes and judicious use of healthcare resources.

390 Disclosures

391 The authors have no financial disclosures.

Funding/Support: HM is funded by a Moorfields Eye Charity PhD Studentship (GR001147). NP is funded by a Moorfields Eye Charity Career Development Award (R190031A). Moorfields Eye Charity is supported in part by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology. ST, BA and DG acknowledge that a proportion of their financial support is from the Department of Health through the award made by the National Institute for Health Research to Moorfields Eye Hospital NHS Foundation Trust and University College London Institute of Ophthalmology for a Specialist Biomedical Research Centre for Ophthalmology. The sponsor or funding organization had no role in the design or conduct of this research. Other acknowledgements: none. References

408		
409	1.	Gordon MO, Steger-May K, Szczotka-Flynn L, et al. Baseline factors
410		predictive of incident penetrating keratoplasty in keratoconus. <i>Am J</i>
411		<i>Ophthalmol.</i> 2006;142(6):923-930.
412	2.	Wittig-Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibson GR. A
413		randomized, controlled trial of corneal collagen cross-linking in progressive
414		keratoconus: three-year results. Ophthalmology. 2014;121(4):812-821.
415	3.	Caporossi A, Mazzotta C, Baiocchi S, Caporossi T. Long-term results of
416		riboflavin ultraviolet A corneal collagen cross-linking for keratoconus in Italy
417		the Siena eye cross study. Am J Ophthalmol. 2010;149(4):585-593.
418	4.	O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised,
419		prospective study to investigate the efficacy of riboflavin/ultraviolet A (370
420		nm) corneal collagen cross-linkage to halt the progression of keratoconus.
421		Br J Ophthalmol. 2011;95(11):1519-1524.
422	5.	Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal
423		crosslinking. J Cataract Refract Surg. 2009;35(8):1358-1362.
424	6.	Gore DM, Leucci MT, Koay SY, et al. Accelerated Pulsed High-Fluence
425		Corneal Cross-Linking for Progressive Keratoconus. Am J Ophthalmol.
426		2021;221:9-16.
427	7.	Salmon HA, Chalk D, Stein K, Frost NA. Cost effectiveness of collagen
428		crosslinking for progressive keratoconus in the UK NHS. Eye .
429		2015;29(11):1504-1511.
430	8.	Lindstrom RL, Berdahl JP, Donnenfeld ED, et al. Corneal cross-linking
431		versus conventional management for keratoconus: a lifetime economic
432		model. J Med Econ. Published online November 19, 2020:1.
433	9.	Godefrooij DA, Mangen MJJ, Chan E, et al. Cost-Effectiveness Analysis of
434		Corneal Collagen Crosslinking for Progressive Keratoconus.
		Ophthalmology. 2017;124(10):1485-1495.

	436	10	Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking
1	437	10.	versus Standard Care on Keratoconus Progression in Young Patients: The
2 3			
4 5	438		KERALINK Randomized Controlled Trial. <i>Ophthalmology</i> . Published online
6	439		April 20, 2021. doi:10.1016/j.ophtha.2021.04.019
7 8	440	11.	Vinciguerra R, Belin MW, Borgia A, et al. Evaluating keratoconus
9 10	441		progression prior to crosslinking: maximum keratometry vs the ABCD
11 12	442		grading system. J Cataract Refract Surg. 2021;47(1):33-39.
13 14			
15	443	12.	Shajari M, Steinwender G, Herrmann K, et al. Evaluation of keratoconus
16 17	444		progression. Br J Ophthalmol. 2019;103(4):551-557.
18 19	445	13.	Ozalp O, Atalay E. Belin ABCD Progression Display Identifies Keratoconus
20 21	446		Progression Earlier Than Conventional Metrics. Am J Ophthalmol.
22 23	447		2022;236:45-52.
24 25	/		
26	448	14.	Hashemi H, Panahi P, Asgari S, Emamian MH, Mehravaran S, Fotouhi A.
27 28	449		Best Indicators for Detecting Keratoconus Progression in Children: A Report
29 30	450		From the Shahroud Schoolchildren Eye Cohort Study. Cornea.
31 32	451		2022;41(4):450-455.
33			
34 35	452	15.	Flynn TH, Sharma DP, Bunce C, Wilkins MR. Differential precision of
36 37	453		corneal Pentacam HR measurements in early and advanced keratoconus.
38 39	454		Br J Ophthalmol. 2016;100(9):1183-1187.
40	155	10	Flackerri F. Llöfner I. Ventheneuleu K. et al. Deliability enalysis of
41 42	455	16.	Flockerzi E, Häfner L, Xanthopoulou K, et al. Reliability analysis of
43 44	456		successive Corneal Visualization Scheimpflug Technology measurements in
45 46	457		different keratoconus stages. Acta Ophthalmol. 2022;100(1):e83-e90.
47 48	458	17.	Kreps EO, Jimenez-Garcia M, Issarti I, Claerhout I, Koppen C, Rozema JJ.
49	459		Repeatability of the Pentacam HR in Various Grades of Keratoconus. Am J
50 51	460		Ophthalmol. 2020;219:154-162.
52 53			
54 55	461	18.	Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part IV:
56	462		further concepts and methods in survival analysis. Br J Cancer.
57 58	463		2003;89(5):781-786.
59 60			
61 62			
63			
64 65			

	464	19.	Kato N, Negishi K, Sakai C, Tsubota K. Baseline factors predicting the need
1 2	465		for corneal crosslinking in patients with keratoconus. PLoS One.
3 4 5	466		2020;15(4):e0231439.
6 7	467	20.	Patrick Royston PL. Flexible Parametric Survival Analysis Using Stata:
8 9	468		Beyond the Cox Model. Stata Press; 2011.
10 11	469	21.	Hardcastle AJ, Liskova P, Bykhovskaya Y, et al. A multi-ethnic genome-wide
12 13	470		association study implicates collagen matrix integrity and cell differentiation
14 15 16	471		pathways in keratoconus. <i>Commun Biol</i> . 2021;4(1):266.
17 18	472	22.	Ding X, Guo X. A Survey of SNP Data Analysis. Big Data Mining and
19 20	473		Analytics. 2018;1(3):173-190.
21 22 23	474	23.	Royston P, Parmar MKB. Flexible parametric proportional-hazards and
24	475		proportional-odds models for censored survival data, with application to
25 26	476		prognostic modelling and estimation of treatment effects. Stat Med.
27 28 29	477		2002;21(15):2175-2197.
30 31	478	24.	Baade PD, Royston P, Youl PH, Weinstock MA, Geller A, Aitken JF.
32 33	479		Prognostic survival model for people diagnosed with invasive cutaneous
34 35	480		melanoma. BMC Cancer. 2015;15:27.
36 37 38	481	25.	Royston P, Parmar MKB, Sylvester R. Construction and validation of a
39	482		prognostic model across several studies, with an application in superficial
40 41	483		bladder cancer: CONSTRUCTION AND VALIDATION OF PROGNOSTIC
42 43 44	484		MODEL. Stat Med. 2004;23(6):907-926.
45 46	485	26.	Royston P, Sauerbrei W. A new measure of prognostic separation in
47 48	486		survival data. Stat Med. 2004;23(5):723-748.
49 50 51	487	27.	Tuft SJ, Moodaley LC, Gregory WM, Davison CR, Buckley RJ. Prognostic
52	488		factors for the progression of keratoconus. Ophthalmology.
53 54 55 56 57 58	489		1994;101(3):439-447.
59 60 61 62 63			
64 65			

490 491 492	28.	Quartilho A, Gore DM, Bunce C, Tuft SJ. Royston–Parmar flexible parametric survival model to predict the probability of keratoconus progression to corneal transplantation. <i>Eye</i> . 2020;34(4):657-662.
493 494 495	29.	Kato N, Masumoto H, Tanabe M, et al. Predicting Keratoconus Progression and Need for Corneal Crosslinking Using Deep Learning. <i>J Clin Med Res</i> . 2021;10(4). doi:10.3390/jcm10040844
496 497	30.	Pearson AR, Soneji B, Sarvananthan N. Does ethnic origin. <i>Eye</i> . 2000;14:625-628.
498 499 500	31.	Georgiou T, Funnell CL, Cassels-Brown A, O'Conor R. Influence of ethnic origin on the incidence of keratoconus and associated atopic disease in Asians and white patients. <i>Eye</i> . 2004;18(4):379-383.
501 502 503 504 505	32.	Royston P. Multivariable Model-Building: A Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Continuous Variables. John Wiley; 2008.

507 Legend

 Figure 1: Chart showing how the Royston-Parmar model fits the entire dataset.
We split the eyes into 4 risk groups by their prognostic index: <25th centile (low risk), 25-50th centile (medium-low risk), 50-75th centile (medium-high risk), >75th centile (high risk). The number of eyes at risk corresponds to the Kaplan-Meier curves.

Figure 2: Time-to-event curves that predict the risk of progression to CXL for three hypothetical patient profiles. The blue line represents a high risk patient who has a 95% probability of progressing to CXL at 5 years. The red line is a medium risk patient who has a 48% probability of progressing to CXL at 5 years. The green line is a low risk patient who has a 14% probability of progressing to CXL at 5 years. The equation used to generate the curves is: $S(t) = e^{-H(t)}$, where H(t) is the cumulative hazard function and is commonly expressed as $ln(H(t)) = s(ln(t)) + x\beta$, where s(ln(t)) is a restricted cubic spline function of log time, β is the vector of coefficients and x is the vector of covariates. For further details of the derivation, we refer the reader to 20. Abbreviations: pachy, pachymetry **Figure 3:** Predicted and observed survival curves for seven postal code regions of Greater London as shown in Supplemental Figure 1 (available at http://www.ajo.com) using IECV. We split the eyes into 4 risk groups by their prognostic index: <25th centile (low risk), 25-50th centile (medium-low risk), 50-75th centile (medium-high risk), >75th centile (high risk). Abbreviations:

- 534 AIC: Akaike Information Criterion
- 58
 59
 535
 535
 535
 536
 537
 538
 538
 539
 535
 535
 536
 537
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538
 538

	536	Back K2: Steep posterior keratometry in the central 3 mm zone
1 2	537	BIC: Bayes Information Criterion
3 4	538	CXL: Corneal Cross-Linking
5	539	EKC: Early Keratoconus Clinic
7	540	EPR: Electronic Patient Record
8 9	541	Front K1: Flat anterior keratometry in the central 3 mm zone
10 11	542	Front K2: Steep anterior keratometry in the central 3 mm zone
12 13	543	HR: Hazard Ratio
14 15	544	IECV: Internal-external Cross Validation
16 17	545	Kmax: Maximum anterior keratometry
18	546	MFP: Multivariate Fractional Polynomial
19 20	547	SNP: Single-nucleotide Polymorphism
21 22	548	
23 24	549	
25 26	550	
27 28	551	
29 30	552	
31 32	553	
33	554	
34 35		
36 37		
38 39		
40 41		
42 43		
44		
45 46		
47 48		
49		
50 51		
52 53		
54 55		
56		
57 58		
59 60		
61 62		
63 64		
65		

1			Style Definition: Normal: Font: (Default) Times Normal:
2		\mathbb{N}	Roman
	-		Style Definition: Header
3	-		Style Definition: Footer
4	-		Formatted: Left: 1.38", Header distance from edge Footer distance from edge: 0.49", Numbering: Cont
5	Personalized model to predict keratoconus progression from		Formatted: Line spacing: 1.5 lines
6	demographic, topographic and genetic data	_	Formatted: Font: Times New Roman, Not Bold, Fo
7			Auto
7 8			
9			
10			Formatted: Line spacing: 1.5 lines
			Formatted: Font: Times New Roman, Not Bold, Fo
11	Short title: Keratoconus progression		Auto
12			Formatted: Line spacing: 1.5 lines
13			Formatted: Font: 9.5 pt
14 15			Formatted: Font: 9.5 pt
	-		Formatted: Font: 9.5 pt
16	-		Formatted: Font: 9.5 pt
17	-		Formatted: Font: 9.5 pt
18	Howard P Maile¹; Ji-Peng Olivia Li₂; Mary D Fortune³; Patrick Royston₄;		Formatted: Font: 9.5 pt
19	Marcello T Leucci ² ; Ismail Moghul ¹ ; Anita Szabo ¹ ; Konstantinos Balaskas ² ;		Formatted: Font: 9.5 pt
			Formatted: Font: 9.5 pt
20	Bruce D Allan ² ; Alison J Hardcastle ¹ ; Pirro Hysi ^{5,6} ; Nikolas Pontikos ¹ *; Stephen J	\langle	Formatted: Font: 9.5 pt
21	Tuft ^{2*} ; Daniel M Gore ^{2*}	\backslash	Formatted: Font: 9.5 pt Formatted: Font: 9.5 pt
22		\mathbb{N}^{\prime}	Formatted: Font: 9.5 pt
23			Formatted: Font: 9.5 pt
24	-		Formatted: Font: 9.5 pt
25	_		Formatted: Font: 9.5 pt
26	UCL Institute of Ophthalmology, 11-43 Bath Street, London EC1V 9EL		Formatted: Font: Times New Roman
		\frown	Formatted: Font: 9.5 pt
27	² Moorfields Eye Hospital NHS Foundation Trust, 162 City Road, London EC1V	$\langle \rangle$	Formatted: Line spacing: 1.5 lines
28	2PD_	$(\)$	Formatted: Font: Times New Roman
29	³ MRC Biostatistics Unit, Cambridge Institute of Public Health, University of	\backslash	Formatted: Font: 9.5 pt
30	Cambridge, UK	$\overline{}$	Formatted: Font: Times New Roman
31	⁴ MRC Clinical Trials Unit at UCL, Aviation House, 125 Kingsway, London,		Formatted: Font: 9.5 pt
			Formatted: Font: Times New Roman
32	WC2B 6NH		Formatted: Font: 9.5 pt Formatted: Font: Times New Roman
33	Section of Ophthalmology, School of Life Course Sciences, King's College		Formatted: Font: 1 miles New Koman
34	London		Formatted: Font: Times New Roman
35	Department of Twin Research and Genetic Epidemiology, King's College		Formatted: Font: 9.5 pt
			Formatted: Font: Times New Roman, Font color: A
36	London		Formatted: Font: +Body (Calibri), Font color: Auto
	2		-
	۲		
	<u> </u>		

37	<u>ــــــــــــــــــــــــــــــــــــ</u>		Formatted: Font: Times New Roman
38	* Authors contributed equally		Formatted: Font: Times New Roman
39	Corresponding author: Mr Daniel Gore, Moorfields Eye Hospital, 162 City		Formatted: Font: 9.5 pt
40	Road, London, EC1V 2PD, UK, <u>daniel.gore1@nhs.net,daniel.gore1@nhs.net.</u>		
41	0207 253 3411.		Formatted: Font: Times New Roman
42	•	•	Formatted: Space After: 0 pt, Line spacing: 1.5 lines
43	Keywords: Keratoconus, corneal cross-linking, keratoconus genetics,		Formatted: Font: Times New Roman
_			Formatted: Line spacing: 1.5 lines
44	keratoconus prediction		Formatted: Font: Times New Roman
45			

2

Formatted: Font: +Body (Calibri), Font color: Auto

Page Break

Keratoconus is a common corneal ectasia that causes irregular astigmatism,

1
2

Introduction

scarring and loss of vision. Thinning and steepening can progress through childhood and early adulthood, but the shape of most eyes stabilizes by the third or fourth decade. Without intervention, keratoconus can lead to severe visual loss, with approximately 10% of eyes eventually requiring corneal transplantation.⁴-Corneal collagen¹ Corneal cross-linking (CXL) by topical application of riboflavin, followed by irradiation with UV-A light, can arrest progression of keratoconus in up to 9888% to 100% of eyes even when there is relatively advanced disease.²⁻⁸²⁶ The potential benefit of CXL is to prevent visual deterioration with a relatively low risk procedure that is cost effective for healthcare providers.⁹⁻¹¹²⁻⁹ However, CXL is usually not offered to all patients at presentation because the disease may have already stabilized. In the recent KERALINK study 43% of children <17 years of age at presentation had not progressed after 18 months.⁸ The definition of progression also varies with the severity of keratoconus, but for early disease a common threshold is either an increase in the maximum keratometry (Kmax) of >1 dioptre, a change in the manifest refractive spherical equivalent of >0.50 dioptre, or an increase in manifest refractive cylinder of >1 dioptre.^{2,12211} Depending on the rate of progression this threshold may be passed in a few months, years, or not at all. At the first assessment it can be a challenge to distinguish eyes that are at risk of rapid progression from those where it is safe to monitor. Unnecessary review visits are a burden to the patient and the care system. We considered the date of numeric progression, $\frac{7_{e}}{2}$ as well as the date when CXL was performed, as alternative end-points to define keratoconus progression. Although the use of keratometry as an end-point may appear the more objective method, there is no consensus variability on the definition of progression reported in the literature and conclusions may vary with the definition that is adopted.^{12,13}¹¹⁻¹⁴ Repeatability thresholds are not usually tailored to individual eyes (i.e. an increase in Kmax by 1 D is not significant in all eyes) and other relevant patient-specific factors are excluded although there is

Formatted: Font: Times New Roman
Formatted: Line spacing: 1.5 lines

Formatted: Font: Times New Roman, Font color: Auto, Pattern: Clear Formatted: Font: Times New Roman

Formatted: Font color: Custom Color(RGB(32,33,34)) Formatted: Font color: Custom Color(RGB(32,33,34))

Formatted: Font: +Body (Calibri), Font color: Auto

1 2				
⊿ 3				
4				
5 6				
7	80	growing evidence on the variability of measurements in more advanced disease		
8	81	and the need for tailoring numerical progression definitions to the disease state,		
9 10	82	and distinguishing real progression from inherent variability of measurement		
11	83	modalities. ¹⁴¹ ₁₆₋₁₇ Finally, patients who receive CXL prior to progression must be		
12	84	censored from the dataset even though these eyes are likely to have been at		
13 14	85	risk of progression. This type of informative censoring creates a bias. ⁴⁵		
15	86	contrast, the time to CXL depends on several variables that include numeric		
16 17	80 87	disease progression, but also incorporates patient-specific risk factors for future		
18	88	progression. Its strength is that it is an easily comprehensible and meaningful		
19				
20 21	89 00	end-point for patients. It encompasses individual risk factors that are not		
22	90	considered when imaging is used in isolation and it has been used by others as	5	
23	91	defining the event of interest. ¹⁶		Formatted: Font: Times New Roman
24 25	92	•		Formatted: Font: Times New Roman
26	93	For these reasons we have used demographic and serial tomography data from		
27	94	a large cohort of patients to generate a time-to-event model to predict the		
28 29	95	probability of an individual progressing to CXL. Because the Cox proportional		
30	96	hazards method does not generate smooth time-to-event curves, we used the		
31	97	Royston-Parmar model to achieve direction estimates of the hazard function. ⁴⁷ ²⁰		
32 33	98	We also performed a further analysis of a subset of patients who had genetic		
34	99	data in the form of single-nucleotide polymorphisms (SNP) generated as part of		
35	100	a study to determine keratoconus risk. ¹⁸ 21	< >	Formatted: Pattern: Clear Formatted: Font: Times New Roman
36 37	101	•	<u> </u>	Formatted: Font: Times New Roman
38	102	Methods		Formatted: Font: Times New Roman
39 40	103	Cohort	1	Formatted: Font: Times New Roman
40 41	104	The study protocol was reviewed and approved by the Clinical Audit		
42	105	Assessment Committee of Moorfields Eye Hospital NHS Foundation Trust		
43 44	106	(reference CA17/CED/03). Institutional Review Board (IRB) approval was		
45	107	obtained and individual patient consent was not required. The study conformed		
46	108	to the tenets of the Declaration of Helsinki. We identified from the Moorfields		
47 48	109	Eye Hospital electronic health record database (OpenEyes) patients aged 13		
49	110	years and above diagnosed with clinical or suspected keratoconus who		
50	111	attended our Early Keratoconus Clinic (EKC) between January 2011 and		
51 52	112	November 2020. Clinical data included keratometry (Kmax, Front K1, Front K2,		
53	113	Back K1, Back K2), and pachymetry (minimum corneal thickness) captured by		Formatted: Font: +Body (Calibri), Font color: Auto
54		2		
55 56		<u>۸</u>		
57				
58 59				
59				
60				
60 61				
60 61 62				
60 61				

1		
4 5		
6		
-	114	Scheimpflug tomography (Pentacam HR, Oculus GmbH, Wetzlar). We only
3	115	included scans with a quality score of 'good' or 'ok', and where multiple scans
	116	were taken on the same day we used the mean value. The date of all CXL
	117	procedures was recorded. The protocol for offering CXL throughout the study
	118	period was, i) a documented history prior to referral to the EKC of our hospital
	119	of significant recent disease progression, \mathcal{F}_{E} ii) a change in contemporary
	120	measurements of 95% above the repeatability limits of the baseline
	121	measurements as shown in Supplemental Table 1 (available at
	122	http://www.ajo.com), $^{7_{E}}$ or iii) a patient considered by a clinician to be at high risk
	123	of progression despite their not fulfilling the above two criteria. Exclusion criteria
	124	included pregnancy or breastfeeding, uncontrolled ocular surface disease or a
	125	minimum corneal thickness less than 375 µm.
	126	
	120	
	127	-
	128	All the data used for model fitting started from the first appointment in the EKC. \leftarrow
	129	Patient demographics included age, gender, smoking status (current or ex/non-
	130	smoker) ethnicity and postcode. Ethnicity was coded as 1 for 'Black' or 'South
	131	Asian or South Asian British' and 0 for any other category (excluding missing
	132	values). Before model fitting, the pachymetry in microns was divided by 10 to
	133	generate a meaningful scale. For the primary analysis, eyes with any missing
	134	data were excluded. We also explored multiple imputation, which avoids data
	135	exclusion by generating multiple versions of the dataset with missing values
	136	replaced with values sampled from an appropriate distribution. To see whether
	137	genetic data can help predict keratoconus progression, we used 28 candidate
	138 hao	SNPs from a recent keratoconus genome-wide association study that contained
	139	926 patients from Moorfields Eye Hospital. ¹⁸ ²¹ The SNP data was encoded as
	140	either 0 (homozygous reference genotype), 1 (heterozygous genotype), or 2
	141	(homozygous variant genotype). We chose to use an additive encoding, thus
	142	the risk of disease increases additively with the degree of genetic variation. ⁴⁹ $_{\mathbb{Z}}$
	143	Anonymized data were then exported to Excel software for analysis (version
	144	15.24 2016, Microsoft Corp.).
	145	•
	146	Model Fitting and Covariate Selection
	ĺ	2
		·/

65

Formatted: Font: Times New Roman, Font color: Auto, Pattern: Clear

Formatted: Line spacing: 1.5 lines

Formatted: Font: Times New Roman Formatted: Font: Times New Roman Formatted: None, Line spacing: 1.5 lines Formatted: Font: Times New Roman, Not Bold Formatted: Font: +Body (Calibri), Font color: Auto

 Formatted: Line spacing: 1.5 lines

Formatted: Font: Times New Roman						
Formatted: Font: Times New Roman						
Formatted: None, Line spacing: 1.5 lines						
Formatted: Font: Times New Roman, Not Bold						
Formatted: Line spacing: 1.5 lines						

-{	Formatted: Font: Times New Roman					
-{	Formatted: Font: Times New Roman					
-(Formatted: None, Line spacing: 1.5 lines					
1	Formatted: Font: Times New Roman, Not Bold					
1	Formatted: Line spacing: 1.5 lines					

Formatted: Font: +Body (Calibri), Font color: Auto

1 2					
3					
4 5					
6					
7 8	181	both the model fitted from data excluding region k ($D_{\underline{k}^{(k)}}$) and also the model	(Formatted: Font: 9.5 pt	
8 9	182	applied to region k (D_k) . ²³ The difference between these two discrimination	-(Formatted: Font: 9.5 pt	
LO	183	metrics $(D_{k}-D_{k})$ was calculated with its corresponding standard error to assess	$\leq \geq$	Formatted: Font: 9.5 pt	
L1 L2	184	the predictive ability of the model. To demonstrate how the model could be	!	Formatted: Font: 9.5 pt	
.3	185	used in practice, we include three hypothetical patients' eyes with different			
4	186	progression risk profiles (high, medium, low risk) and plot the predicted time-to-			
.5 .6	187	event curve for each shown in Figure 2.	(Formatted: Font: Times New Roman	
.7	188	<u>ــــــــــــــــــــــــــــــــــــ</u>	-(Formatted: Font: Times New Roman	
.8 .9	189	Statistical Analysis	$\sim >$	Formatted: None, Line spacing: 1.5 lines	
9 0	190	The event of interest was defined as the date that the eye underwent CXL. We \checkmark	$\sim >$	Formatted: Font: Times New Roman, Not Bold	
1	191	calculated the time-to-event as the difference between the first appointment in	Ŀ	Formatted: Line spacing: 1.5 lines	
2 3	192	our service and the date of CXL (or the last patient appointment in the case of			
4	193	censoring). Since we had paired observations (eyes), we used variance-			
5	194	corrected models to account for correlation between eyes and to ensure that			
6 7	195	robust standard errors were produced. The choice of scale and selection of			
8	196	degrees of freedom for the Royston-Parmar model was informed by inspecting			
9 0	197	the Akaike information criterion (AIC) and Bayes information criterion (BIC) $^{\tt 17_{22}}$			
1	198	and the results of this were balanced with ease of interpretation. See			
2	199	Supplemental Table 2 and Supplemental Text 1 (available at at			
3 1	200	http://www.ajo.com) for further explanation. Royston and Sauerbrei's D statistic			
)	201	was used as a measure of discrimination and $R_{_{2D}}^{2}$ -as a measure of explained	-(Formatted: Font: 9.5 pt	
5	202	variation (both calculated on the natural scale of the model). Although all of the			
	203	primary results were generated from a complete case analysis, we performed			
)	204	an additional analysis using multiple chained imputation (predictive mean			
	205	matching approach with 5 nearest neighbors). Model fitting was performed in			
	206	Stata 13 (StataCorp LP, Texas, USA) and the Royston-Parmar model was fitted			
3 1	207	using the stpm2 package from Stata 13.	-(*	Formatted: Font: Times New Roman	
	208	•	-(*	Formatted: Font: Times New Roman	
,	209	Results	(Formatted: Font: Times New Roman	
	210	Cohort	\frown	Formatted: Font: Times New Roman, Not Bold	
	211	From a potential of 9,341 eyes (4316 pairs of eyes and 709 individual eyes), the⊷	$\sim >$	Formatted: None, Line spacing: 1.5 lines Formatted: Line spacing: 1.5 lines	—
	212	final model used 8,701 eyes of 4,823 patients, with 3,232 eyes that had CXL.		Formaticu: Line spacing. 1.5 mes	
2	213	The mean age was 28.3 years with standard deviation of 7.1 years. We			
3	214	excluded 640 eyes with missing data. Table 1 summarizes the available	ſ	Formatted: Font: +Body (Calibri), Font color: Auto	
:		2			
5		•			
7					
8 9					
0					
1					

covariates along with missing data percentages. See Supplemental Text 2 and
Supplemental Table 3 (available at-<u>http://www.ajo.com</u>) for a description of the
multiple imputation results.

219 Model Fitting and Covariate Selection (Genetic Data)

We analyzed patients with genetic data separately because this data was only available for ~14% of patients. Of 926 patients (1852 eyes) with genetic data, 531 eyes were excluded with incomplete keratometry or CXL data, which left 1321 eyes, of which 665 had CXL. With univariate analysis of the 28 SNPs only rs72631889 was found to be significant (P=0.01) (Supplemental Table 4 (available at at http://www.ajo.com)). We then produced a multivariable model via backwards selection on this subset of eyes using corneal data, patient data and rs72631889 as an additional covariate as shown in Supplemental Table 5 (available at-_http://www.ajo.com). However rs72631889, although significant (P=0.005), had a negligible contribution (0.3%) to the explained variation in the final model.

232 Model Fitting and Covariate Selection (Excluding Genetic Data)

The results of the univariate time-to-event analysis on the hazards scale using a Royston-Parmar flexible parametric model is shown in Table 2. Genetic data was excluded from this analysis. All variables except smoking status were significant. The explained variation (R_{2D}^{2}) and discrimination (D) were highest for age (17%) and Kmax (15%) with Front K1, Front K2, Back K1, Back K2 and pachymetry each explaining 6-10% of the variation. Notably, gender and ethnicity, although significant in the univariate analysis, did not contribute to explained variation. The hazard ratios for significant covariates indicate that increasing age at presentation, greater pachymetry and flatter (less negative) posterior keratometry values decrease risk of having CXL, whilst steeper anterior keratometry values and male gender increase the risk of having CXL.

When we fitted a multivariable model the significant covariates were age,
Kmax, Front K1, Front K2 and pachymetry (Table 2). When we removed single
variables from the model the effect this had on explained variation and

-{	Formatted: Pattern: Clear
-{	Formatted: Font: Times New Roman
-(Formatted: Font: Times New Roman
-(Formatted: Font: Times New Roman

-{	Formatted: Font: Times New Roman
-(Formatted: Font: Times New Roman
-(Formatted: None, Line spacing: 1.5 lines
\neg	Formatted: Font: Times New Roman, Not Bold
\neg	Formatted: Line spacing: 1.5 lines
1	Formatted: Font: Times New Roman

Formatted: Font: 9.5 pt

-	Formatted: Pattern: Clear
1	Formatted: Font: Times New Roman
-(Formatted: Font: Times New Roman

Formatted: Font: +Body (Calibri), Font color: Auto

3			
4			
5			
6 7	249	discrimination is shown in Supplemental Table 6 (available at at	
8 9	250	http://www.ajo.com). Age was the most important covariate (16.7%), with Kmax	
10	251	contributing ~5% of explained variation. K1, K2 and pachymetry had a small	
11	252	effect (<1%) when removed individually. We chose a model without K2 on the	
12 13	253	basis of parsimony, which was supported by the fact that K1 and K2 were	
14	255	highly correlated (R_{2}^{2} =0.91) as shown in Supplemental Figure 2 (available at at	
15	255	http://www.ajo.com). The final fitted model hazard ratios can be seen on the	
16	255 256		
17 18		multivariable column of Table 2. It is notable that an increase in K1 now has a	
19	257	protective effect in the final model. The explained variation and discrimination	
20	258	for the final model were 32.7% and 1.43 respectively. ²⁴ / ₂₂ The opposing effect of	
21 22	259	Kmax and Front K1 can be explained by examining their regression coefficients	
23	260	before converting to hazard ratios; Kmax has a positive coefficient (0.0795) and	
24	261	Front K1 has a negative coefficient (-0.0749). This is logically similar to	
25 26	262	including the combined covariate (Kmax - Front K1) in the model which can be	
20 27	263	viewed clinically as a proxy for irregular astigmatism. We also investigated	
28	264	combining K1 and K2 into a single covariate as K2-K1 (standard definition of	
29	265	astigmatism), but the corresponding p value was not significant.	
30 31	266		
32	267	Figure 1 visually depicts the result of applying the final model to the original	
33	268	dataset. As expected, the predicted mean survival curves closely follow the	
34 35	269	Kaplan-Meier curves. To demonstrate the use of the model in clinical practice,	
36	270	survival curves for three hypothetical patients followed for five years are shown	
37	271	in Figure 2. We have also produced a free web application from the model	
38 39	Г 272	which can be accessed at	
40	272	https://pontikoslab.com/kcprog.http://beta.moorfieldscxl.com.	
41	273		
42 43	274 275	Keratometric Progression Sensitivity Analysis	
44			
45	276	The results of the keratometric progression sensitivity analysis can be found in	-
46 47	277	the Supplementary Material. By examining the Kaplan Meier curves in	
48	278	Supplemental Figure 3, we can see that the best case time-to-event curve	
49	279	indicates a 40% survival probability at 5 years whilst the worst case curve	
50	280	indicates a 27% survival probability at 5 years. This 13% differecedifference in	
51 52	281	survival probability at 5 years represents the upper bound of the discrepancy in	
53	282	survival probability within the data. After fitting the Royston-Parmar model,	
54		2	/
55 56		· /	
57			
58			
59 60			
00			

Formatted: Font: 9.5 pt

 Formatted: Pattern: Clear

 Formatted: Font: Times New Roman

 Formatted: Pattern: Clear

 Formatted: Font: Times New Roman

 Formatted: Font: Times New Roman

 Formatted: Font: Times New Roman

 Formatted: None, Line spacing: 1.5 lines

 Formatted: Font: Times New Roman, Not Bold

 Formatted: Line spacing: 1.5 lines

Formatted: Font: +Body (Calibri), Font color: Auto

283	amongst the bazard ratios which everlap (ago, kmaxk/max, k2), there was			
	amongst the hazard ratios which overlap (age, <u>kmaxKmax</u> , k2), there was			
284 be5	reasonable similarity (Supplemental Tables 8 and 9). Most importantly, the			
285	model fitted to the best case data had an explained variation of 11% compared			
286	to 23% for the worst case indicating a significant difference in model			
287 bee	performance depending on the assumptions used for handling eyes which			
288	received CXL.			Formatted: Font: Times New Roman, Font color: Aut Pattern: Clear
289	• Multivariable Madel Velidetien			Formatted: Font: Times New Roman
290	Multivariable Model Validation		_	Formatted: None, Line spacing: 1.5 lines Formatted: Font: Times New Roman, Not Bold
291	When performing validation using internal-external cross validation, Figure 3	•		Formatted: Line spacing: 1.5 lines
292	shows the ability of our final model to predict keratoconus progression across			
293	different geographic regions. We did not identify any significant differences in			
294	prognostic factors across regions. The model prediction curves generally follow			
295	the Kaplan Maier curves. Notably, region 5 (South West Greater London) and			
296	region 7 (other regions) have a worse predictive performance than the other			
297	regions, indicating that these regions have different characteristics compared			
298	with the remainder of the dataset used for model fitting. This could be due to			
299	differing patient characteristics, such as complex cases that required referral to			
300	our tertiary referral centercentre rather than being managed locally. Overall, the			
301	prediction becomes less accurate over time, which is expected due to low			
302	numbers with follow-up beyond three years. Supplemental Table 7 (available			
303	at http://www.ajo.com displays quantitative validation results of the model			
304	using internal external validation. The difference column $D_{\underline{k}} - D_{\underline{k}}$ is a measure of		\leq	Formatted: Font: 9.5 pt Formatted: Font: 9.5 pt
305	predictive ability. Region 7 (other regions outside of Greater London) has the			Formatted: Font. 9.5 pt
306	greatest discrepancy in discrimination (-0.26) which indicates that the model			
307	fitted when excluding region 7 had greater discriminative ability than when			
308	applied to region 7 alone.		_	Formatted: Font color: Custom Color(RGB(32,33,34) Pattern: Clear (White)
309	+			Formatted: Font: Times New Roman
310	Discussion			Formatted: Font: Times New Roman
311	▲			Formatted: Font: Times New Roman
312	In this study we have incorporated demographic, keratometric, and genetic data			
313	to generate a prognostic model of keratoconus progression to CXL. We have			
314	shown that parameters recorded at the first examination (age, Kmax, Front K1,			
315	minimum pachymetry) can produce a time-to-event curve to calculate a			
316	personalized risk for keratoconus progression. Although we chose time to CXL		/	Formatted: Font: +Body (Calibri), Font color: Auto
	2	/		
1	۸			

317	rather than keratometric progression as the end point for the time-to-event	
318	analysis, we performed a sensitivity analysis using keratometric progression,	
319	and found that a CXL model accounts for a much higher proportion of the	
320	explained variation (33%) compared to the keratometric model (11% or 23% for	
320	best and worst case respectively). The opposing effects of Kmax and Front K1	
321	were surprisingunexpected, but similar to including the combined covariate	
323	(Kmax - Front K1) in the model; a possible explanation is that the opposing	
323	effect is the result of an increase in irregular astigmatism. Of the significant	Formatted: Font color: Text 1
325	covariates in our model, younger age made the greatest contribution to our	
326	model. Thus, one should have a lower threshold for treatment in younger	
327	patients.	
328	When applying internal-external cross validation, the survival curves closely	Formatted: Line spacing: 1.5 lines
329	followed the Kaplan Meier survival curves for each of the geographic regions,	Formatted. Elle spacing. 1.5 mes
329	which indicates generalisability, and model discrimination between training and	
331		
332	cross validation groups was similar, indicating that the predictive ability is well	
	maintained. Finally, our SNP genetic data had limited additional predictive utility	
333	for keratoconus progression. However, the genetic dataset was relatively small	
334	(926 patients), and recruitment was based on the presence of keratoconus, as	
335	opposed to the severity of keratoconus, or any other index of risk of rapid	
336	progression.	Formatted: Font: Times New Roman, Font color: Auto, Pattern: Clear
337		Formatted: Font: Times New Roman, Font color: Auto, Pattern: Clear
338	The Royston-Parmar model has previously been used to predict the likelihood	
339	of the worst eye of patients with keratoconus progressing to corneal	
340	transplantation. ²⁵ ²⁶ In their final model, Quartilho et al chose 3 significant	
341	covariates: Kmax, age and ethnicity. The reported covariate hazard ratios that	
342	overlap with our study (Kmax and age) were different in magnitude but in the	
343	same direction. When performing internal validation their model exhibited good	
344	predictive ability. They produced time-dependent receiver operating curves	
345	using the validation set and found one-year sensitivity and specificity to be	
346	92.8% and 94.6% respectively. Using logistic regression, Kato et al. found that	
347	the two strongest factors associated with the requirement for CXL were age and	
348	Kmax, which is consistent with our findings. ^{46_19} Moreover the team went on to	
349	find that age combined with corneal tomography maps was able to predict	Formatted: Font: Times New Roman
350	progression and need for crosslinking using deep learning.	Formatted: Font: +Body (Calibri), Font color: Auto
	2	
I	٠	

351	•	Formatted: Font: Times New Roman
352	An ability to generate personalized time-to-event curves that predict	
353	progression to CXL (Figure 2) could directly inform clinical decisions that benefit	
354	patient care. Firstly, patients may better understand their own risk for	
355	progression and feel more confident in choosing their management options.	
356	Secondly, for both clinicians and patients, the prediction of progression may	
357	contribute to scheduling treatments, including prioritizing patients at high risk of	
358	early progression. For example, patients at high risk with a 98% probability of	
359	progressing to CXL at 5 years could be offered CXL at the point of first	
360	diagnosis without waiting to demonstrate keratometric progression. Medium risk	
361	patients may benefit from a period of clinician-led topographic monitoring. For	
362	the lowest risk patients, optometry-led monitoring in the community may be	
363	sufficient. This risk stratification could be tailored to regions and reflect local	
364	needs and resources such as provision of monitoring services in regions with	
365	lower risk and greater capacity for CXL in areas with more high risk patients.	
366	Finally, when a decision is made to postpone CXL for further monitoring, the	
367	time-to-event curve can contribute to decisions on the scheduling of future	
368	follow up reviews, with perhaps shorter time periods where the curve is	
369	steepestRecommendations based on this model on clinical practice is yet to	
370	be evaluated.	Formatted: Font: Times New Roman
371	•	Formatted: Font: Times New Roman
372	Our study is subject to several limitations inherent to our dataset. First, if	
373	patients had CXL at another hospital, this may not be reliably recorded in the	
374	source database. This could lead to a very small number of patients being	
375	included in the analysis who have already had CXL. Second, ethnicity is a well	
376	knownestablished risk factor for keratoconus and keratoconus	
377	progression, ^{24,26,27} 27.30.31 but ethnicity is now an optional field at patient registration	
378	at our institution and this information was unavailable for approximately 50% of	
379	our dataset. However, even when we restricted the dataset to those with	
380	ethnicity records, it was not found to be a significant covariate. Third, though	
381	the cohort used for univariable and multivariable analysis were identical the	
382	number of eyes where all covariates were available was lower than for	
383	univariable analysis due to missing data. Finally, when we used multiple	
384	imputation to generate a multivariable model, ethnicity was still not found to be	Formatted: Font: +Body (Calibri), Font color: Auto
1	2	
	•/	

385	significant. In the model fitting process we chose to use a simple backwards		
386	selection as opposed to the multivariate fractional polynomial (MFP) method. ²⁸ $_{28}$		
387	In our initial investigations, the results of MFP yielded nonlinear functional forms		
388	of the covariates and, whilst this method may have slightly increased the		
389	predictive power of the prognostic model, the resulting hazard ratios would be		
390	very hard to interpret. In addition, we did not examine time dependent effects		
391	for the covariates, which may provide a more accurate model fit, and future		
392	studies should examine this option. Finally, although no external validation		
393	dataset was available, internal external cross validation allowed us to confirm		
394	that our model is generalizable across geographical regions.	Fo	ormatted: Font: Times New Roman
395			ormatted: Font: Times New Roman
396	In conclusion, we have fitted a prognostic model for progression of keratoconus		
397	to CXL which generates a time-to-event curve using age, Kmax, Front K1,		
398	minimum pachymetry from time of presentation. Incorporation of a relatively		
399	small genetic dataset does not improve the explained variation of our model.		
400	Personalized modeling of risk may improve patients' understanding of their		
401	condition and the need for CXL. Such a model may help better improve patients		
402	and aid clinician decision making to CXL to achieve better outcomes and		
403	judicious use of healthcare resources.	Fo	ormatted: Font: Times New Roman, Font color: Auto
404	·		ormatted: Font: Times New Roman, Font color: Auto
405	Disclosures	Fo	ormatted: Font: Times New Roman, Not Bold, Font col
406	The authors have no financial disclosures.		uto prmatted: Font: Times New Roman
407	A		ormatted: Font: Times New Roman
408	Funding/Support: HM is funded by a Moorfields Eye Charity PhD Studentship		
409	(GR001147). NP is funded by a Moorfields Eye Charity Career Development		
410	Award (R190031A). Moorfields Eye Charity is supported in part by the National		
411	Institute for Health Research (NIHR) Biomedical Research Centre based at		
412	Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of		
413	Ophthalmology. ST, BA and DG acknowledge that a proportion of their financial		
414	support is from the Department of Health through the award made by the		
415	National Institute for Health Research to Moorfields Eye Hospital NHS		
416	Foundation Trust and University College London Institute of Ophthalmology for		
417	a Specialist Biomedical Research Centre for Ophthalmology. The sponsor or	Fo	ormatted: Font: Times New Roman
418	funding organization had no role in the design or conduct of this research.		prmatted: Font: +Body (Calibri), Font color: Auto
	2		
	•/		

19				Formatted: Font: Times New Roman
20	O	ther acknowledgements: none.		Formatted: Font: Times New Roman
21				
22				
23	-			
24		References_		
25		<u>ــــــــــــــــــــــــــــــــــــ</u>	\succ	Formatted: Font: Times New Roman
26	1.	Gordon MO, Steger-May K, Szczotka-Flynn L, et al. Baseline factors		Formatted: Indent: Left: 0.24", Line spacing: 1.5 lines
27		predictive of incident penetrating keratoplasty in keratoconus. Am J		
28		Ophthalmol. 2006;142(6):923-930.		
29	2.	Wittig-Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibson GR. A		
30		randomized, controlled trial of corneal collagen cross-linking in progressive		
31		keratoconus: three-year results. Ophthalmology. 2014;121(4):812-821.		
32	3.	Caporossi A, Mazzotta C, Baiocchi S, Caporossi T. Long-term results of		
33	0.	riboflavin ultraviolet A corneal collagen cross-linking for keratoconus in Italy:		
34		the Siena eye cross study. <i>Am J Ophthalmol.</i> 2010;149(4):585-593.		
35	4.	<u>O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised,</u>		
36		prospective study to investigate the efficacy of riboflavin/ultraviolet A (370		
37		nm) corneal collagen cross-linkage to halt the progression of keratoconus.		
38		Br J Ophthalmol. 2011;95(11):1519-1524.		
39	5.	_Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal		Formatted: Indent: Left: -0.1", Hanging: 0.33", Space
40		crosslinking. J Cataract Refract Surg. 2009;35(8):1358-1362.		After: 12 pt, Line spacing: 1.5 lines
		·····		Formatted: Font: Times New Roman
41	6.	Gore DM, Leucci MT, Koay SY, et al. Accelerated Pulsed High-Fluence		
42		Corneal Cross-Linking for Progressive Keratoconus. Am J Ophthalmol.		
43		2021;221:9-16.		Formatted: Font: Times New Roman
44	7.	Salmon HA, Chalk D, Stein K, Frost NA. Cost effectiveness of collagen		
45		crosslinking for progressive keratoconus in the UK NHS. Eye .		
46		2015;29(11):1504-1511,	_	Formatted: Font: Times New Roman
				Formatted: Font: +Body (Calibri), Font color: Auto
		2		
	.			

 After: 12 pt, Line spacing: 1.5 lines Formatted: Font: Times New Roman 	147 <u>8</u>	Lindstrom RL, Berdahl JP, Donnenfeld ED, et al. Corneal cross-linking	
 Source Standard Care on Kerateconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. 2017;124(10):1485-1495, Source Standard Care on Kerateconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. 2017;124(10):1485-1495, Source Standard Care on Kerateconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. 2017;124(10):1485-1495, Source Standard Care on Kerateconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. 2014;12(4):812-821. Coprosci A, Mazzutta C, Baisechi S, Caporessi T, Long Form results of ribollavin ultraviolet A cormaal collagen cross-linking for Kerateconus. Brogression of Kerateconus in taly: the Siena -voo cross-etudy. Am J Ophthalmol. 2010;140(1):812-821. Coprosci A, Mazzutta C, Baisechi S, Caporessi T, Long Form results of ribollavin ultraviolet A cormaal collagen cross-linking for Kerateconus. Brogression of Kerateconus Stream Parking of Progression Kerateconus. Am J Ophthalmol. 2010; Stream Stream Parking Versus Standard Care on Kerateconus. Am J Ophthalmol. 2010; Stream Versus S	48	versus conventional management for keratoconus: a lifetime economic	
S1 Corneal Collagen Crosslinking for Progressive Keratoconus. S2 Ophthalmology. 2017;124(10):1485-1495, S3 Carkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking S4 versus Standard Care on Keratoconus Progression in Young Patients: The S4 KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:4. Gorden MO, Steger May K, Szezetka Flynn L, et al. S5 Bacchic factocop predictive of incident peneturating keratoplasty in Koratoconus. S6 April 20, 2021. doi:4. S6 Corneal Collagen cross-linking in progressive Keratoconus. Keratoconus: three year results. Ophthalmology. 2014;121(4):812-821. S6 Coperossi A, Mazzetta C, Baiecchi S, Caporossi T. Long term results of tabolavin ultraviolet A corneal collagen cross-linking for keratoconus. <i>Brill</i> S6 Ophthalmol. 2014;132(4):485-603. S70 2013;27(3):320-339. S71 S6 S6 Gore DM, Shortt AJ, Allan BD. New elinical pathways for keratoconus. <i>Eye</i> . S72 Versute: Four: Time New Roman S73 2013;27(3):320-339. S74 Gore DM, Shortt AJ, Allan BD. New elinical pathways for keratoconus. <i>Eye</i> . S75 Core DM, Loucei MT, Koay S Y, et al. Accelorated	149	model. J Med Econ. Published online November 19, 2020:1.	
 5.1 Corneal Collagen Crossiluking for Progressive Keratoconus. 5.2 Ophthalmology. 2017;124(10):1485-1495, 5.3 Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking 5.4 versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online 5.6 April 20, 2021, doi:4. Corden Hoo, Steger May K, Szezetka Flynn L, et al. 5.8 Baceline factors predictive of incident peortexing kerateplasty in koratoconus. 5.9 Wittig Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibeon GR. A 5.1 randomized, controlled trial of control of the peortexing in progressive keratoconus: three-yoar results. Ophthalmol. 2010;140(1):855-693. 5.6 O'Brart DPS, Chan E, Samarae K, Patel P, Shah SP. A randomised, prospective study. Am J Ophthalmol. 2010;140(1):455-693. 5.6 Gore DM, Shortt AJ, Alian BD. New clinical pathways-for keratoconus. Br J 6. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal 7.1 Gore DM, Leucci MT, Koay S Y, et al. Accelerated Pulsed High Fluence 7.2 Gore DM, Leucci MT, Koay S Y, et al. Accelerated Pulsed High Fluence 7.2 Gore DM, Leucci MT, Koay S Y, et al. Accelerated Pulsed High Fluence 7.4 Gore DM, Leucci MT, Koay S Y, et al. Accelerated Pulsed High Fluence 7.5 2021;221:0:16₂, 10:1016/j.ophtha.2021.04.019, 10:	50 <u>9</u>	0Godefrooij DA, Mangen MJJ, Chan E, et al. Cost-Effectiveness Analysis of 🛛 🖛	
 Sato, Larkin DEP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology</i>. Published online Am J Ophthalmol. 2006;142(6):923-930. Writig Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibson GR. A randomized, controlled trial of corneal collagen cross-linking in progressive keratosonus: three year results. <i>Ophthalmology</i>. 2014;121(4):812-821. Caperossi A, Mazzotta C, Baioschi S, Caporossi T. Long term results of ribollayin ultraviolet A corneal collagen cross-linking for keratosonus. Intaly: the Siena eye cross study. <i>Am J Ophthalmol.</i> 2010;140(4):585-593. O' Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised. prospective study to investigate the efficacy of ribollayin/ultraviolet A (370 nm) corneal collagen cross-linkage to halt the progression of keratosonus. <i>Er J</i> <i>Qohthalmol.</i> 2011;92(11):1610-1624. Gendere DM, Shortt AJ, Allan BD. New clinical pathways for keratosonus. <i>Er J</i> 2013;27(3):329-339. Koller T, Mrochen M, Seiler T, Complication and failure rates after corneal crosslinking. <i>J Cataract Refract Surg.</i> 2009;35(8):1358-1382. Formatted: Poir: Times New Roman KeRALINK Randomized Controlled Trial. <i>Ophthalmology</i>. Published online April 20, 2024. doi:10.1016/j.ophtha.2021.04.019. Formatted: Poir: Times New Roman 	51	Corneal Collagen Crosslinking for Progressive Keratoconus.	After: 12 pt, Line spacing: 1.5 lines
 Versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:1. Gero DM, Scieger May K, Szezetka Flynn L, et al. Baseline factors predictive of incident penetrating keratoplasty in keratoconus. Am J Ophthalmol. 2006;142(6):923-930. Wittig Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibson GR. A randomized, controlled trial of corneal collagen cross-linking in progressive keratoconus: three year results. Ophthalmology. 2014;121(4):812-821. Caporossi A, Mazzetta C, Baioschi 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;149(4):585-593. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised. prospective study to investigate the officacy of iboflavin/ultraviolet A (370 nm) corneal collagen cross-linkage to halt the progression of keratoconus. <i>Br J</i> Ophthalmol. 2011;95(11):1519-1524. Goro DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. <i>Br J</i> 2013;27(3):329-339. Koller T, Mrochen M, Seller T. Complication and failure rates after corneal corneal Cross Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2024;221:9-16. Core DM, Loucei MT, Koay S Y, et al. Accelerated Pulsed High Fluence Corneal Cross Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2024;221:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology</i>. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Fermatted: Font: Time: New Roman Permatted: Font: Standard Care on Keratoconu	52	<u>Ophthalmology. 2017;124(10):1485-1495.</u>	Formatted: Font: Times New Roman
 KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:4Gordon MO, Steger May K, Szczetka Flynn L, et al. Beceline factore predictive of incident ponetrating koratoplasty in koratoconus. <i>Am J Ophthalmol.</i> 2006;142(6):023-030. Wittig Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibeon GR. A randomized, controlled trial of corneal collagen cross-linking in progressive keratoconus: three-year results. <i>Ophthalmology</i>. 2014;121(4):812-821. Caporossi A, Mazzetta C, Baioschi S, Caporossi T. Long torm results of ribollavin ultraviolet A conneal collagen cross-linking for koratoconus in flaty: the Siena eye cross citudy. <i>Am J Ophthalmol.</i> 2011;19(1):49(4):685-503. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the officacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross-linking to progression of keratoconus. <i>Br J</i> <i>Ophthalmol.</i> 2011;195(11):4519-1524. Gere DM, Shortt AJ, Allan BD. New elinical pathways for keratoconus. <i>Br J</i> 2013;27(3):320-339. Formattel: Foit: Times New Roman Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal eroesclinking. <i>J Catarast: Refract Surg.</i> 2000;35(8):1358-1362. Formattel: Foit: Times New Roman Corneal Cross- Linking for Progressive Koratoconus. <i>Am J Ophthalmol.</i> 2021;221:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross- linking versus Standard Care on Keratoconus. <i>Part J Ophthalmol.</i> 2021;221:9-16. Formattel: Foit: Times New Roman KERALINK Randomized Controlled Trial. <i>Ophthalmology</i>. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formattel: Foit: Times New Roman Formattel: Foit: Times New Roman Formattel: Foit: Times New Roman 			
 April 20, 2021. doi:4Cordon MO, Steger May K, Szezetka Flynn L, et al. Baceline factors predictive of incident ponetrating keratoplasty in keratoconus. Am J Ophthalmol. 2006;142(6):023-030. Wittig Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibson GR. A randomized, controlled trial of corneal collagen cross-linking in progressive keratoconus: three year results. Ophthalmology. 2014;121(4):812-821. Caporossi A, Mazzotta C, Baiocchi S, Caporossi T. Long term results of ribollavin ultraviolet A corneal collagen cross-linking for keratoconus. In taly: the Siena eye cross study. Am J Ophthalmol. 2010;140(4):585-503. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the efficacy of ribollavin/ultraviolet A (370 nm) corneal collagen cross-linking to healt the progression of keratoconus. Br J Ophthalmol. 2011;95(11):1519-1524. Core DM, Shortt AJ, Allan BD. New elinical pathways for keratoconus. Br J 2013;27(3):320-330, Formattet: Index: Left: 0.1°, Hanging: 0.33°, Spa Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. J Cataract Refract Surg. 2000;35(8):1358-1362, Formattet: Foir: Times New Roman Corneal Cross- Linking for Progressive Keratoconus. Am J Ophthalmol. 2021;221:9-16, Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross- linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Centrolled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formattet: Foir: Times New Roman Formattet: Foir: Times New Roman Formattet: Foir: Times New Roman 			
 Baseline factors predictive of incident penetrating keratoplasty in keratoconus. <i>Am J Ophthalmol.</i> 2006;142(6):923-930. Wittig Silva C, Chan E, Islam FMA, Wu T, Whiting M, Snibson GR. A randomized, controlled trial of corneal collagen cross-linking in progressive keratoconus: three-year results. <i>Ophthalmology</i>. 2014;121(4):812-821. Caporossi A, Mazzotta C, Baiocchi S, Caporossi T. Long-term results of iboflavin ultraviolet A corneal collagen cross-linking for keratoconus in Italy: the Siona eye cross study. <i>Am J Ophthalmol.</i> 2010;14(4):585-593. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the officacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross-linking to halt the progression of keratoconus. <i>Br J Ophthalmol.</i> 2011;195(11):1519-1524. Gore DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. <i>Eye</i>. • 2013;27(3):329-339. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. <i>J Cataraet Rofraet Surg.</i> 2009;35(8):1358-1362. Formatted: Foit: Times New Roman KeRALINK Randomized Controlled Trial. <i>Ophthalmol.</i> 201: 40:19 Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology</i>. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Foit: Times New Roman 			
 Wittig Silva C, Chan E, Helam FMA, Wu T, Whiting M, Snibson GR. A randomized, controlled trial of corneal collagen cross-linking in progressive keratecenus: three-year results. <i>Ophthalmol.</i>2914;121(4):812-821. Caporesi A, Mazzotta C, Baioschi S, Caporesci T, Long torm rosults of ribollavin ultraviolet A corneal collagen cross-linking for keratecenus: in Italy: the Siena eye cross study. <i>Am J Ophthalmol.</i> 2011;149(4):865-593. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the efficacy of ribollavin/ultraviolet A (370 nm) corneal collagen cross-linking for keratecenus. <i>Br J Ophthalmol.</i> 2011;195(11):1610-1524. Gere DM, Shortt AJ, Allan BD. New clinical pathways for keratecenus. <i>Eye</i>. Kotler T, Mrochon M, Seller T. Complication and failure rates after corneal cross-linking. <i>J Cataract Refract Surg.</i> 2009;36(8):1358-1362. Kotler T, Mrochon M, Seller T. Complication and failure rates after corneal cross-linking for Progressive Keratecenus. <i>Am J Ophthalmol.</i> 2021;221:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratecenus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmol.</i>997. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman Formatted: Font: Times New Roman 			
 fandomized, controlled trial of correal collagen cross-linking in progressive kerateconus: three year results. <i>Ophthalmology</i>. 2014;121(4):812-821. Caporossi A, Mazzotta C, Baiocchi S, Caporossi T. Long-term results of ritoldivin ultraviolet A correal collagen cross-linking for kerateconus in Italy: the Siena eye cross study. <i>Am J Ophthalmol</i>. 2010;149(4):585-593. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the officacy of riboflavin/ultraviolet A (370 nm) correal-collagen cross-linking for herateconus. <i>Br J Ophthalmol</i>. 2011;96(11):4510-4524. Gere DM, Shortt AJ, Allan BD. New clinical pathways for kerateconus. <i>Eye</i>. Koller T, Mrochen M, Seiler T. Complication and failure rates after correal cross-linking for Progressive Kerateconus. <i>Am J Ophthalmol</i>. 2021;221:9-16. Keration DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Kerateconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology</i>. Publiched online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman Formatted: Font: Times New Roman Formatted: Font: Times New Roman 	158	Am J Ophthalmol. 2006;142(6):923-930.	
 kerateconus: three year results. Ophthalmology. 2014;121(4):812-521. Caporossi A, Mazzotta C, Baiocchi S, Caporossi T. Long term results of riboflavin ultraviolet A corneal collagen cross-linking for kerateconus in Italy: the Siena ovo cross study. Am J Ophthalmol. 2010;140(4):585-593. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, propective study to investigate the officacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross-linkage to hall the progression of kerateconus. Br J Ophthalmol. 2011;140(1):585-593. Ger Gore DM, Shortt AJ, Allan BD. New clinical pathways for kerateconus. Br J Ophthalmol. 2013;27(3):329-339. Ger Gore DM, Shortt AJ, Allan BD. New clinical pathways for kerateconus. Eye., * 2013;27(3):329-339. Corneal Cross-linking of T. Complication and failure rates after corneal crosselinking. J Cataraet Refract Surg. 2009;35(8):1358-1362. Formatted: Font: Times New Roman Corneal Cross-Linking for Progressive Kerateconus. Am J Ophthalmol. 2021;221:9-16. Corneal Cross-Linking for Progressive Kerateconus. Am J Ophthalmol. 2022;221:21:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Kerateconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 			
 Caporossi A, Mazzuta C, Baiocchi S, Caporossi T. Long torm results of ribollavin ultraviolet A corneal collagen cross-linking for keratoconus in Italy: the Siona eye cross study. <i>Am J Ophthalmol.</i> 2010;149(4):585-593. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the efficacy of ribollavin/ultraviolet A (370 nm) corneal collagen cross-linking to halt the progression of keratoconus. <i>Br J Ophthalmol.</i> 2011;9(11):1610-1624. Ger Gore DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. <i>Eye</i>, * 2013;27(3):329-339. Formatted: Inden: Left: -0.1", Hanging: 0.33", Spa After: 12 pt, Line spacing: 1.5 lines Portiated: Font: Times New Roman Keller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosse linking. <i>J Cataract Refract Surg.</i> 2009;36(8):1358-1362. Formatted: Font: Times New Roman Corneal Cross Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:0-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross- linking versue Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology.</i> Published online April-20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 			
 ribółlavin ultraviolet A corneal collagen cross-linking for Keratoconus in Italy: the Siona cyc cross study. Am J Ophthalmol. 2010;149(4):585-593. O'Brart DPS, Chan E, Samaras K, Patel P, Shah SP. A randomised, prospective study to investigate the officacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross-linking to hab the progression of keratoconus. <i>Br J</i> <i>Ophthalmol.</i> 2011;95(11):1519-1524. Gereo DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. <i>Eye</i>. 2013;27(3):329-330. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. <i>J Cataract Refract Surg.</i> 2009;35(8):1358-1362. Formatted: Font: Times New Roman Corneal Cross-Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:9:16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology.</i> Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 			
 Siona eye cross study. Am J Ophthalmol. 2010;140(4):585-593. C'Brart DPS, Chan E, Samarae K, Patel P, Shah SP, A randomised, prospective study to investigate the officacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross linkage to halt the progression of keratoconus. <i>Br J</i> <i>Ophthalmol.</i> 2011;95(11):1519-1524. Gere DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. <i>Eye.</i> 4 2013;27(3):329-339. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. <i>J Cataract Rofract Surg.</i> 2000;35(8):1358-1362. Gorne DM, Leucci MT, Koay S-Y, et al. Accelerated Pulsed High Fluence Corneal Cross Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:0-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology</i>. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 			
 prospective study to investigate the efficacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross linkage to halt the progression of keratoconus. <i>Br J Ophthalmol.</i> 2011;05(11):1519-1524. Gere DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. <i>Eye.</i> • 2013;27(3):329-339. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal cross-linking. <i>J Cataract Refract Surg.</i> 2009;35(8):1358-1362. Formatted: Font: Times New Roman Corneal Cross-Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology.</i> Published online April-20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 			
 prospective study to investigate the efficacy of riboflavin/ultraviolet A (370 nm) corneal collagen cross linkage to halt the progression of koratoconus. <i>Br J Ophthalmol.</i> 2011;95(11):1519-1524. Gere DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. <i>Eye.</i> • 2013;27(3):329-339. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal cross-linking. <i>J Cataract Refract Surg.</i> 2009;35(8):1358-1362. Formatted: Font: Times New Roman Corneal Cross-Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology.</i> Published online April-20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 	651	O'Brart DPS Chan E. Samaras K. Datel P. Shah SP. A randomised	
 Ophthalmol. 2011;95(11):1519-1524. Gere DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. Eye. 2013;27(3):329-339, Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. <i>J Cataract Refract Surg.</i> 2009;35(8):1358-1362, Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal cross-linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:9-16, Corneal Cross-Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:9-16, Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology.</i> Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 			
 Gore DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. Eye., * 2013;27(3):329-339. Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. J Cataract Refract Surg. 2009;35(8):1358-1362. Formatted: Font: Times New Roman Gore DM, Leucci MT, Koay S-Y, et al. Accelerated Pulsed High Fluence Corneal Cross-Linking for Progressive Keratoconus. Am J Ophthalmol. 2021;221:9-16. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 			
 After: 12 pt, Line spacing: 1.5 lines Formatted: Font: Times New Roman Koller T, Mrochen M, Seiler T. Complication and failure rates after corneal crosslinking. <i>J Cataract Refract Surg.</i> 2009;35(8):1358-1362. Formatted: Font: Times New Roman Formatted: Font: Times New Roman Corneal Cross-Linking for Progressive Keratoconus. <i>Am J Ophthalmol.</i> 2021;221:9-16. Se Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. <i>Ophthalmology.</i> Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. Formatted: Font: Times New Roman 	68	Ophthalmol. 2011;95(11):1519-1524.	
 2013;27(3):329-339, Formatted: Font: Times New Roman 	69 5	. —Gore DM, Shortt AJ, Allan BD. New clinical pathways for keratoconus. Eye . ←	
 472 crosslinking. J Cataract Refract Surg. 2009;35(8):1358-1362. 473 7. Gore DM, Leucci MT, Koay S-Y, et al. Accelerated Pulsed High Fluence Corneal Cross Linking for Progressive Keratoconus. Am J Ophthalmol. 2021;221:9-16. 475 2021;221:9-16. 476 8. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. 476 Formatted: Font: Times New Roman 477 Formatted: Font: Times New Roman 478 Formatted: Font: Times New Roman 479 April 20, 2021. doi:10.1016/j.ophtha.2021.04.019. 	170	2013;27(3):329-339,	
472 crosslinking. J Cataract Refract Surg. 2009;35(8):1358-1362. Formatted: Font: Times New Roman 473 7. Gore DM, Leucci MT, Koay S-Y, et al. Accelerated Pulsed High-Fluence Corneal Cross-Linking for Progressive Keratoconus. Am J Ophthalmol. 475 2021;221:9-16. Formatted: Font: Times New Roman 476 8. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The 478 KERALINK Randomized Controlled Trial. Ophthalmology. Published online Formatted: Font: Times New Roman 479 April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman	171 6	Koller T. Mrochen M. Seiler T. Complication and failure rates after corneal	
 473 7. Gore DM, Leucci MT, Koay S-Y, et al. Accelerated Pulsed High Fluence Corneal Cross-Linking for Progressive Keratoconus. Am J Ophthalmol. 2021;221:9-16. 8. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman Formatted: Font: Times New Roman 			Formattade Font: Times Naw Poman
 Corneal Cross-Linking for Progressive Keratoconus. Am J Ophthalmol. 2021;221:9-16, Formatted: Font: Times New Roman 8. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman Formatted: Font: +Body (Calibri), Font color: Auto 	+/2		Formatter, Font. Times New Roman
 2021;221:9-16, Formatted: Font: Times New Roman Earkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019, Formatted: Font: Times New Roman Formatted: Font: Times New Roman 	1 73 ∓		
 476 8. Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking 477 versus Standard Care on Keratoconus Progression in Young Patients: The 478 KERALINK Randomized Controlled Trial. Ophthalmology. Published online 479 April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman Formatted: Font: +Body (Calibri), Font color: Auto 	474	Corneal Cross-Linking for Progressive Keratoconus. Am J Ophthalmol.	
 versus Standard Care on Keratoconus Progression in Young Patients: The KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman Formatted: Font: +Body (Calibri), Font color: Auto 	75	2021;221:9-16.	Formatted: Font: Times New Roman
KERALINK Randomized Controlled Trial. Ophthalmology. Published online April 20, 2021. doi:10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman Formatted: Font: +Body (Calibri), Font color: Auto	76 🔒	Larkin DFP, Chowdhury K, Burr JM, et al. Effect of Corneal Cross-linking	
April 20, 2021. doi: 10.1016/j.ophtha.2021.04.019 Formatted: Font: Times New Roman Formatted: Font: +Body (Calibri), Font color: Auto	77	versus Standard Care on Keratoconus Progression in Young Patients: The	
Formatted: Font: +Body (Calibri), Font color: Auto	178	KERALINK Randomized Controlled Trial. Ophthalmology. Published online	
	479	April 20, 2021. doi: 10.1016/j.ophtha.2021.04.019	Formatted: Font: Times New Roman
			Formattadi Fonti (Pody (Calibri) Font color: Auto
<u><u></u></u>		2	Formatted: Font: +Body (Calibri), Font Color: Auto
		±	

	<u>11.</u>	Vinciguerra R, Belin MW, Borgia A, et al. Evaluating keratoconus	
81		progression prior to crosslinking: maximum keratometry vs the ABCD	
82		grading system. J Cataract Refract Surg. 2021;47(1):33-39.	Formatted: Font: Times New Roman
83	12.	Shajari M, Steinwender G, Herrmann K, et al. Evaluation of keratoconus	
84		progression. Br J Ophthalmol. 2019;103(4):551-557.	Formatted: Font: Times New Roman
85	13.	Ozalp O, Atalay E. Belin ABCD Progression Display Identifies Keratoconus	
86		Progression Earlier Than Conventional Metrics. Am J Ophthalmol.	
87		2022;236:45-52.	Formatted: Font: Times New Roman
88	1/	_Hashemi H, Panahi P, Asgari S, Emamian MH, Mehravaran S, Fotouhi A.	
89	14.	Best Indicators for Detecting Keratoconus Progression in Children: A Report	
90		From the Shahroud Schoolchildren Eye Cohort Study. <i>Cornea</i> .	
91		2022;41(4):450-455.	Formatted: Font: Times New Roman
92	15.	Flynn TH, Sharma DP, Bunce C, Wilkins MR. Differential precision of	
93		corneal Pentacam HR measurements in early and advanced keratoconus.	
94		Br J Ophthalmol. 2016;100(9):1183-1187.	Formatted: Font: Times New Roman
95	16.	Flockerzi E, Häfner L, Xanthopoulou K, et al. Reliability analysis of	
96		successive Corneal Visualization Scheimpflug Technology measurements in	
97		different keratoconus stages. Acta Ophthalmol. 2022;100(1):e83-e90.	Formatted: Font: Times New Roman
.98	17	_Kreps EO, Jimenez-Garcia M, Issarti I, Claerhout I, Koppen C, Rozema JJ.	
.99		Repeatability of the Pentacam HR in Various Grades of Keratoconus. Am J	
00		Ophthalmol. 2020;219:154-162.	Formatted: Font: Times New Roman
	18.	<u>Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part IV:</u>	
02		further concepts and methods in survival analysis. Br J Cancer.	
03		<u>2003;89(5):781-786.</u>	
04	19.	Kato N, Negishi K, Sakai C, Tsubota K. Baseline factors predicting the need	Formatted: Indent: Left: -0.1", Hanging: 0.33", Space After: 12 pt, Line spacing: 1.5 lines
05		for corneal crosslinking in patients with keratoconus. PLoS One.	After: 12 pt, Line spacing: 1.5 lines
06		<u>2020;15(4):e0231439.</u>	Formatted: Font: Times New Roman
			Formatted: Font: +Body (Calibri), Font color: Auto
		2	
	.	/	

08	Beyond the Cox Model. Stata Press; 2011.	Formatted: Font: Times New Roman
09 <u>21.</u>	Hardcastle AJ, Liskova P, Bykhovskaya Y, et al. A multi-ethnic genome-wide	
10	association study implicates collagen matrix integrity and cell differentiation	
511	pathways in keratoconus. Commun Biol. 2021;4(1):266.	Formatted: Font: Times New Roman
12 <u>22.</u>	Ding X, Guo X. A Survey of SNP Data Analysis. Big Data Mining and	
13	Analytics. 2018;1(3):173-190.	
14 <u>23.</u>	Royston P, Parmar MKB. Flexible parametric proportional-hazards and	
15	proportional-odds models for censored survival data, with application to	
16	prognostic modelling and estimation of treatment effects. Stat Med.	
517	<u>2002;21(15):2175-2197.</u>	
18 <u>24.</u>	Baade PD, Royston P, Youl PH, Weinstock MA, Geller A, Aitken JF.	
19	Prognostic survival model for people diagnosed with invasive cutaneous	
20	melanoma. BMC Cancer. 2015;15:27.	
21 <u>25.</u>	Royston P, Parmar MKB, Sylvester R. Construction and validation of a	Formatted: Indent: Left: -0.1", Hanging: 0.33", Space
22	prognostic model across several studies, with an application in superficial	After: 12 pt, Line spacing: 1.5 lines
23	bladder cancer: CONSTRUCTION AND VALIDATION OF PROGNOSTIC	
24	MODEL. Stat Med. 2004;23(6):907-926.	Formatted: Font: Times New Roman
25 <u>26.</u>	Royston P, Sauerbrei W. A new measure of prognostic separation in	
26	survival data. Stat Med. 2004;23(5):723-748.	
27 <u>27.</u>	Tuft SJ, Moodaley LC, Gregory WM, Davison CR, Buckley RJ. Prognostic	
28	factors for the progression of keratoconus. Ophthalmology.	
29	<u>1994;101(3):439-447.</u>	
30 <u>28.</u>	Quartilho A, Gore DM, Bunce C, Tuft SJ. Royston-Parmar flexible	
31	parametric survival model to predict the probability of keratoconus	
32	progression to corneal transplantation. Eye . 2020;34(4):657-662.	
		Terms H. J. Derts D. J. (O.P. N. D. J. 1997)
	2	Formatted: Font: +Body (Calibri), Font color: Auto
.	/	

	00	Keta N. Maanmata I.I. Taraha M. at al. Davidation Kanata annu Davidara	
33 34	<u>29.</u>	Kato N, Masumoto H, Tanabe M, et al. Predicting Keratoconus Progression and Need for Corneal Crosslinking Using Deep Learning. J Clin Med Res.	
35		2021;10(4). doi:10.3390/jcm10040844	
6	<u>30.</u>	Pearson AR, Soneji B, Sarvananthan N. Does ethnic origin. Eye .	
37		2000;14:625-628.	
88	<u>31.</u>	Georgiou T, Funnell CL, Cassels-Brown A, O'Conor R. Influence of ethnic	
9		origin on the incidence of keratoconus and associated atopic disease in	
0		Asians and white patients. Eye . 2004;18(4):379-383.	
1	<u>32.</u>	Royston P. Multivariable Model-Building: A Pragmatic Approach to	
2		Regression Analysis Based on Fractional Polynomials for Continuous	
3		<u>Variables. John Wiley; 2008.</u>	
4	_		
5			
			F
		2	
	.	/	

Formatted: Font: +Body (Calibri), Font color: Auto

	2	Formatted: Font: +Body (Calibri), Font color: Auto
7 21-8 73 74	Hardcastle AJ, Liskova P, Bykhovskaya Y, et al. A multi-ethnic genome-wide association study implicates collagen matrix integrity and cell differentiation pathways in kerateconus. Commun Biol. 2021;4(1):266.	
1	Beyond the Cox Model. Stata Press; 2011.	Formatted: Font: Times New Roman
0	17. Patrick Royston PL. Flexible Parametric Survival Analysis Using Stata:	
69	2020;15(4):e0231439,	Formatted: Font: Times New Roman
57 58	16. Kato N, Negishi K, Sakai C, Tsubota K. Baseline factors predicting the need for corneal crosslinking in patients with keratoconus. PLoS One.	
6	2003;89(5):781-786.	Formatted: Font: Times New Roman
64 65	45.—Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part IV: further concepts and methods in survival analysis. Br J Cancer.	
3	Br J Ophthalmol. 2016;100(9):1183-1187-	Formatted: Font: Times New Roman
51 52	14. Flynn TH, Sharma DP, Bunce C, Wilkins MR. Differential precision of corneal Pentacam HR measurements in early and advanced keratoconus.	
60	progression. Br J Ophthalmol. 2019;103(4):551-557.	Formatted: Font: Times New Roman
9	13. Shajari M, Steinwender G, Herrmann K, et al. Evaluation of keratoconus	
7 8	progression prior to crosslinking: maximum keratometry vs the ABCD grading system. <i>J Cataract Refract Surg.</i> 2021;47(1):33-39.	Formatted: Font: Times New Roman
6	12. Vinciguerra R, Belin MW, Borgia A, et al. Evaluating keratoconus	
5	Ophthalmology. 2017;124(10):1485-1495.	Formatted: Font: Times New Roman
3 4	11.—Godefrooij DA, Mangen M-JJ, Chan E, et al. Cost-Effectiveness Analysis of ← Corneal Collagen Crosslinking for Progressive Keratoconus.	Formatted: Indent: Left: -0.1", Hanging: 0.33", Spac After: 12 pt, Line spacing: 1.5 lines
0 10 1 2	Lindstrom RL, Berdahl JP, Donnenfeld ED, et al. Corneal cross-linking versus conventional management for keratoconus: a lifetime economic model. J Med Econ. Published online November 19, 2020:1.	
.9	2015;29(11):1504-1511.	Formatted: Font: Times New Roman
.7 -8	 Salmon HA, Chalk D, Stein K, Frost NA. Cost effectiveness of collagen crosslinking for progressive keratoconus in the UK NHS. Eye . 	Formatted: Indent: Left: -0.1", Hanging: 0.33", Spac After: 12 pt, Line spacing: 1.5 lines
-6		

	19. – Ding X, Guo X. A Survey of SNP Data Analysis. Big Data Mining and	4	Formatted: Indent: Left: -0.1", Hanging: 0.33", Space After: 12 pt, Line spacing: 1.5 lines
576	Analytics. 2018;1(3):173-190,		Formatted: Font: Times New Roman
577 🕯	20Royston P, Parmar MKB. Flexible parametric proportional-hazards and		
578	proportional-odds models for censored survival data, with application to		
579	prognostic modelling and estimation of treatment effects. Stat Med.		
580	2002;21(15):2175-2197.		Formatted: Font: Times New Roman
	21. Baade PD, Royston P, Youl PH, Weinstock MA, Geller A, Aitken JF.		
582	Prognostic survival model for people diagnosed with invasive cutaneous		
583	melanoma. BMC Cancer. 2015;15:27.		Formatted: Font: Times New Roman
58 422.	-Royston P, Parmar MKB, Sylvester R. Construction and validation of a		
585 586 587	prognostic model across several studies, with an application in superficial bladder cancer: CONSTRUCTION AND VALIDATION OF PROGNOSTIC MODEL. Stat Med. 2004;23(6):907-926.		
58 <mark>&3.</mark>	Royston P, Sauerbrei W. A new measure of prognostic separation in survival		
589	data. Stat Med. 2004;23(5):723-748.		
	Tuft SJ, Moodaley LC, Gregory WM, Davison CR, Buckley RJ. Prognostic		
591 592	factors for the progression of keratoconus. <i>Ophthalmology</i> . 1994;101(3):439-447.		
593	25. Quartilho A, Gore DM, Bunce C, Tuft SJ. Royston-Parmar flexible	4	Formatted: Indent: Left: -0.1", Hanging: 0.33", Space
594	parametric survival model to predict the probability of keratoconus		After: 12 pt, Line spacing: 1.5 lines
595	progression to corneal transplantation. Eye . 2020;34(4):657-662.		Formatted: Font: Times New Roman
59 26. 597	-Pearson AR, Soneji B, Sarvananthan N. Does ethnic origin. <i>Eye</i> . 2000;14:625- 628.		
5007	Georgiou T, Funnell CL, Cassels-Brown A, O'Conor R. Influence of ethnic origin		
59 & 1. 599	on the incidence of keratoconus and associated atopic disease in Asians and	t	
600	white patients. Eye . 2004;18(4):379-383.		
60 <mark>28.</mark>	Royston P. Multivariable Model-Building: A Pragmatic Approach to Regression		
602 603	Analysis Based on Fractional Polynomials for Continuous Variables. John Wiley; 2008.		
	micy, 2000.		
604 605	Legend	•	Formatted: Font color: Black
606	A 34		Formatted: Line spacing: 1.5 lines
607			Formatted: Font: Times New Roman, Not Bold
608	-		
			Formatted: Font: +Body (Calibri), Font color: Auto
	2		
	<u>ـ</u>		

9	Figure 1: Chart showing how the Royston-Parmar model fits the entire dataset.	Formatted: Line spacing: 1.5 lines
0	We split the eyes into 4 risk groups by their prognostic index: <25th centile (low	
1	risk), 25-50th centile (medium-low risk), 50-75th centile (medium-high	
12	risk), >75th centile (high risk). The number of eyes at risk corresponds to the	
13	Kaplan-Meier curves.	Formatted: Font: Times New Roman
14	•	Formatted: Font: Times New Roman
15	Figure 2: Time-to-event curves that predict the risk of progression to CXL for	
16	three hypothetical patient profiles. The blue line represents a high risk patient	
17	who has a 95% probability of progressing to CXL at 5 years. The red line is a	
18	medium risk patient who has a 48% probability of progressing to CXL at 5	
19	years. The green line is a low risk patient who has a 14% probability of	
20	progressing to CXL at 5 years. The equation used to generate the curves is:	
21	$S(t) = e^{-H(t)}$, where $H(t)$ is the cumulative hazard function and is commonly	
22	expressed as $ln(H(t)) = s(ln(t)) + x\beta$, where $s(ln(t))$ is a restricted cubic	
23	spline function of log time, β is the vector of coefficients and x is the vector of	
24	covariates. For further details of the derivation, we refer the reader to	
25	<u>∞.</u> Abbroviatione: pachy, pachymotry.	
26	Abbreviations: pachy, pachymetry	
27	+	Formatted: Line spacing: 1.5 lines
28	Figure 3: Predicted and observed survival curves for seven postal code regions	Formatted: Font: Times New Roman
29	of Greater London as shown in Supplemental Figure 1 (available at	
30	http://www.ajo.com) using IECV. We split the eyes into 4 risk groups by their	
31	prognostic index: <25th centile (low risk), 25-50th centile (medium-low risk), 50-	
32	75th centile (medium-high risk), >75th centile (high risk).	Formatted: Font: Times New Roman
33	•	Formatted: Font: Times New Roman
34	Abbreviations:	Formatted: Font: Times New Roman
35	AIC: Akaike Information Criterion	Formatted: Font: Times New Roman
36	Back K1: Flat posterior keratometry in the central 3 mm zone	Formatted: Font: Times New Roman
37	Back K2: Steep posterior keratometry in the central 3 mm zone	Formatted: Font: Times New Roman
38	BIC: Bayes Information Criterion	Formatted: Font: Times New Roman Formatted: Font: +Body (Calibri), Font color: Auto
	2	Formatten, Font. +Body (Canon), Font colof: Allto

1 2					
3					
4 5					
6					
7 8	639	CXL: Corneal Collagen-Cross-LinkingLinking			Formatted: Font: Times New Roman
9	640	EKC: Early Keratoconus Clinic			Formatted: Font: Times New Roman
10	641	EPR: Electronic Patient Record			Formatted: Font: Times New Roman
11 12	642	Front K1: Flat anterior keratometry in the central 3 mm zone			Formatted: Font: Times New Roman
13	643	Front K2: Steep anterior keratometry in the central 3 mm zone			Formatted: Font: Times New Roman
14 15	644	HR: Hazard Ratio			Formatted: Font: Times New Roman
16	645	IECV: Internal-external Cross Validation			Formatted: Font: Times New Roman
17	646	Kmax: Maximum anterior keratometry			Formatted: Font: Times New Roman
18 19	647	MFP: Multivariate Fractional Polynomial			Formatted: Font: Times New Roman
20	648	SNP: Single-nucleotide Polymorphism			Formatted: Font: Times New Roman
21 22	649	-			
23	650	-			
24	651				
25 26	652				
27	653				
28	654				
29 30	655	A		<	Formatted: Font: Times New Roman
31	I				Formatted: Indent: Left: 0", Hanging: 0.33", Space After: 12 pt, Tab stops: 0.33", Left
32 33					
33 34					
35					
36 37					
38					
39					
40 41					
42					
43					
44 45					
46					
47					
48 49					
50					
51					
52 53					Formatted: Font: +Body (Calibri), Font color: Auto
54		2			
55			/		
56 57					
58					
59 60					
60 61					
62					
63 64					
65					

Covariate	Туре	Mean	SD	N	Missing No. (%)
Front K1 (D)	Numeric	45.31	3.86	8,813	528 (5.7)
Front K2 (D)	Numeric	48.39	4.85	8,839	502 (5.4)
Back K1 (D)	Numeric	-6.53	0.75	7,949	1392 (14.9)
Back K2 (D)	Numeric	-7.23	0.93	8,702	639 (6.8)
Kmax (D)	Numeric	54.14	8.01	8,834	507 (5.4)
Pachymetry (um)	Numeric	462.92	46.15	8,946	395 (4.2)
Age (years)	Numeric	28.28	7.10	9341	0 (0)
Genetic data ^a	Ordinal	N/A	N/A	1141	8020 (85.9)
Self-reported black or asian ethnicity ^b	Categorical (59.9% black or asian)	N/A	N/A	4889	4452 (47.7)
Male gender	Categorical (67% male)	N/A	N/A	9341	0 (0)
Smoker ^c	Categorical (4.5% smoker)	N/A	N/A	9341	0 (0)

Table 1: Summary statistics for the available covariates at the first examination for 9341 eyes

 recorded at first visit.

Abbreviations: Front K1, flattest anterior keratometry; Front K2, steepest anterior keratometry; Back

K1, flattest posterior keratometry; Back K2, steepest posterior keratometry; Kmax: maximum

Keratometry; pachymetry: minimum corneal thickness; SD, Standard deviation; N, number of eyes;

N/A, not applicable.

^aGenetic data comprised of 28 SNPs and was encoded in an additive fashion (0,1,2).

^b1=black or asian, 0=otherwise. ^c0=non-smoker/ex-smoker, 1=current smoker.

Table 2: Univariable and final multivariable model for all considered covariables excluding genetic

 data in the training dataset fitted on the hazards scale with 5 degrees of freedom.

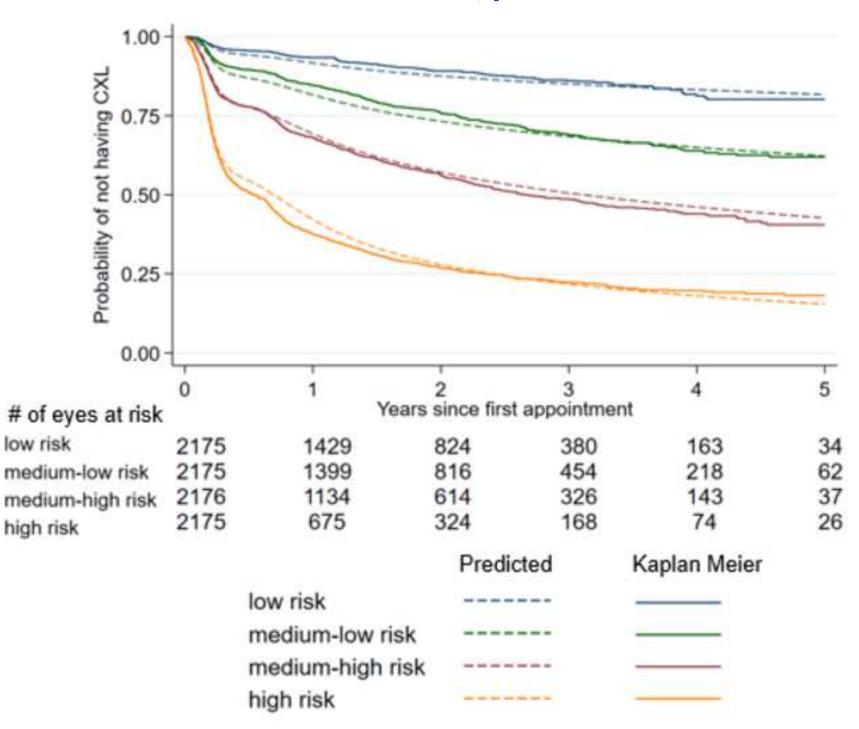
	Univariable (N=9	9341)	Multivariable (N=8701)			
Covariate	Hazard Ratio [95% CI]	P Value	R ² _D	D	Hazard Ratio [95% CI]	P Value
Ethnicity	1.14 [1.02; 1.27]	0.02	0.4%	0.13	N/A	N/A
Smoker ^a	1.07 [0.9; 1.28]	0.46	0.1%	0.05	N/A	N/A
Male Gender	1.11 [1.01; 1.21]	0.02	0.2%	0.10	N/A	N/A
Age at presentation	0.91 [0.9; 0.92]	<0.001	16.7%	0.92	0.9 [0.90; 0.91]	<0.001
Kmax	1.06 [1.05; 1.06]	<0.001	14.9%	0.86	1.08 [1.07; 1.09]	<0.001
Front K1	1.09 [1.08; 1.1]	<0.001	7.0%	0.56	0.93 [0.91; 0.94]	<0.001
Front K2	1.08 [1.07; 1.08]	<0.001	9.8%	0.67	N/A	N/A
Back K1 ^c	0.67 [0.64; 0.71]	<0.001	5.9%	0.51	N/A	N/A
Back K2 ^c	0.7 [0.67; 0.72]	<0.001	8.4%	0.62	N/A	N/A
Pachymetry 10 ^b	0.93 [0.92; 0.94]	<0.001	7.5%	0.58	0.95 [0.93; 0.96]	<0.001

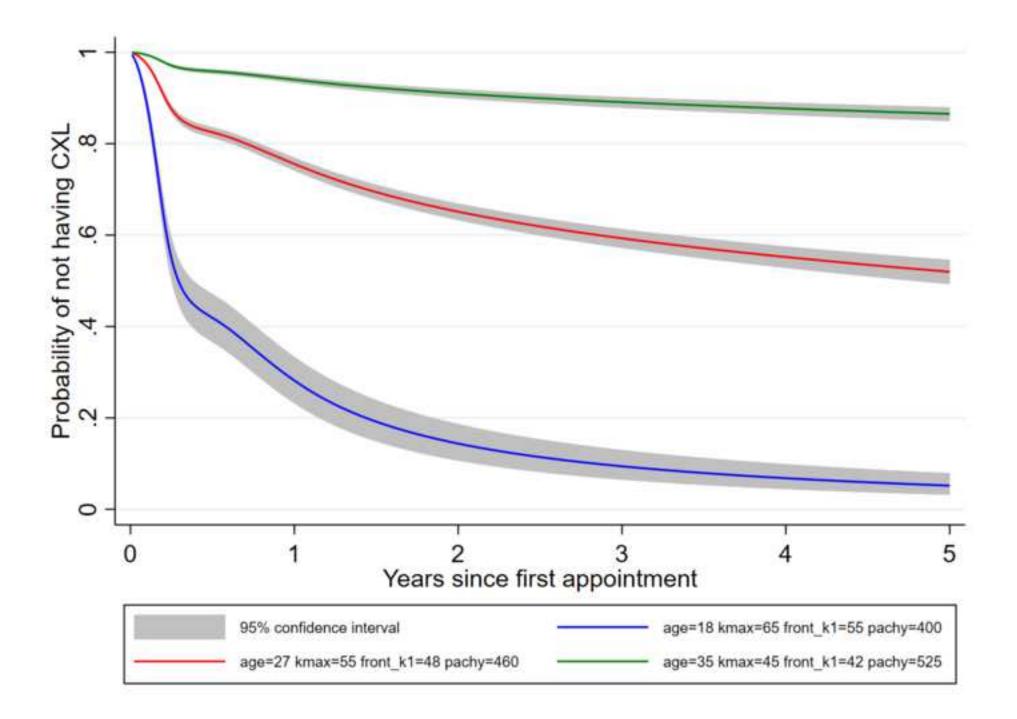
Abbreviations: N, number of eyes; R²_D, explained variation; D, Royston and Sauerbrei's D statistic (used as a measure of discrimination); CI, confidence interval; Kmax, maximum keratometry; Front K1, flattest anterior keratometry; Front K2, steepest anterior keratometry; Back K1, flattest posterior keratometry; Back K2, steepest posterior keratometry; pachymetry, minimum corneal thickness; N/A, not applicable due to this variable not being included in the final model.

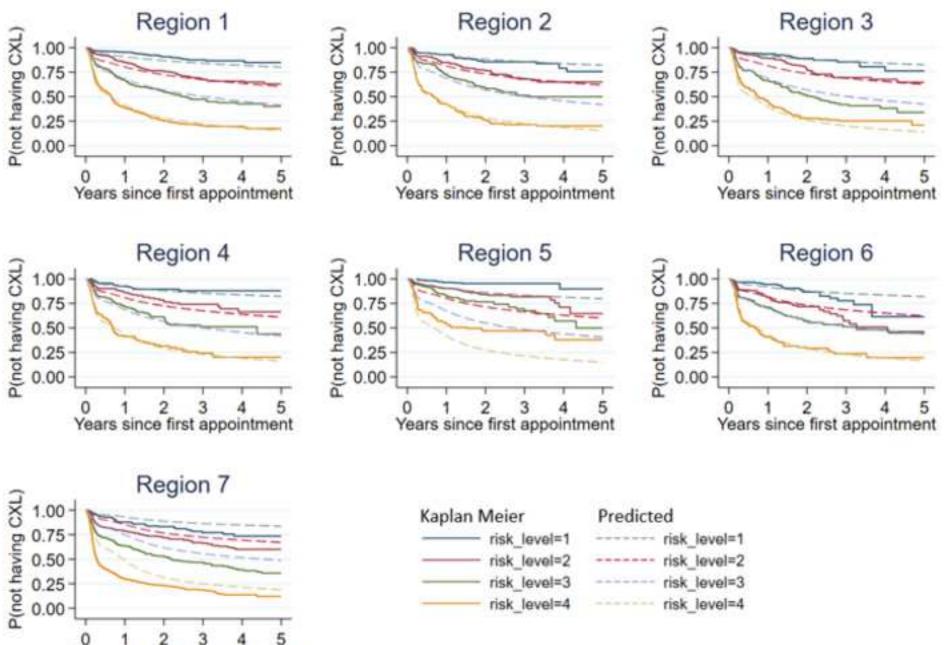
^a0=non-smoker/ex-smoker, 1=current smoker.

^bMinimum pachymetry in steps of 10µm.

^cBack K1 and Back K2 are negative values such that patients with advanced keratoconus are typically associated with large negative values. A hazard ratio below 1 indicates that as measurements become more positive, the risk of progression decreases.

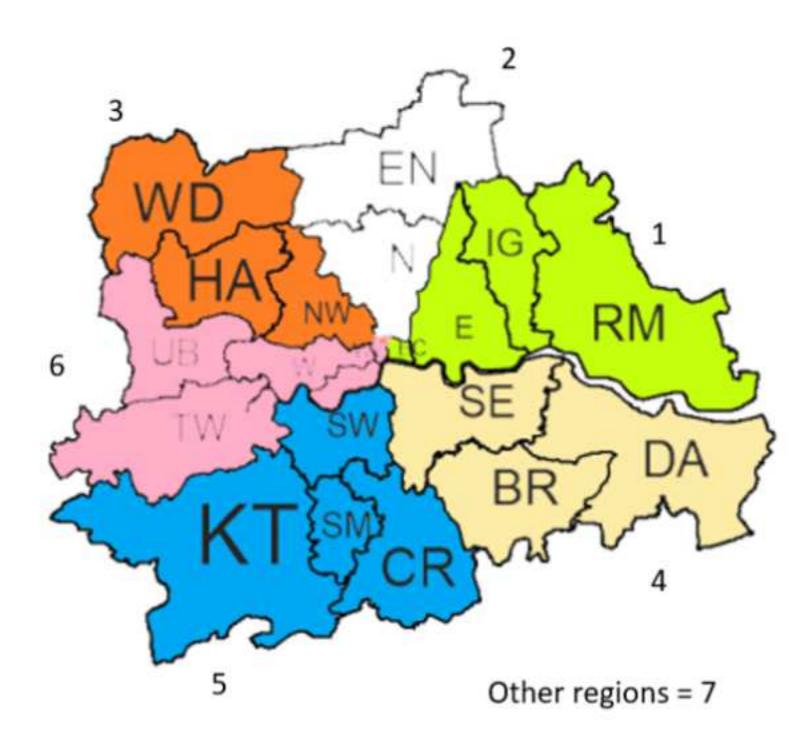






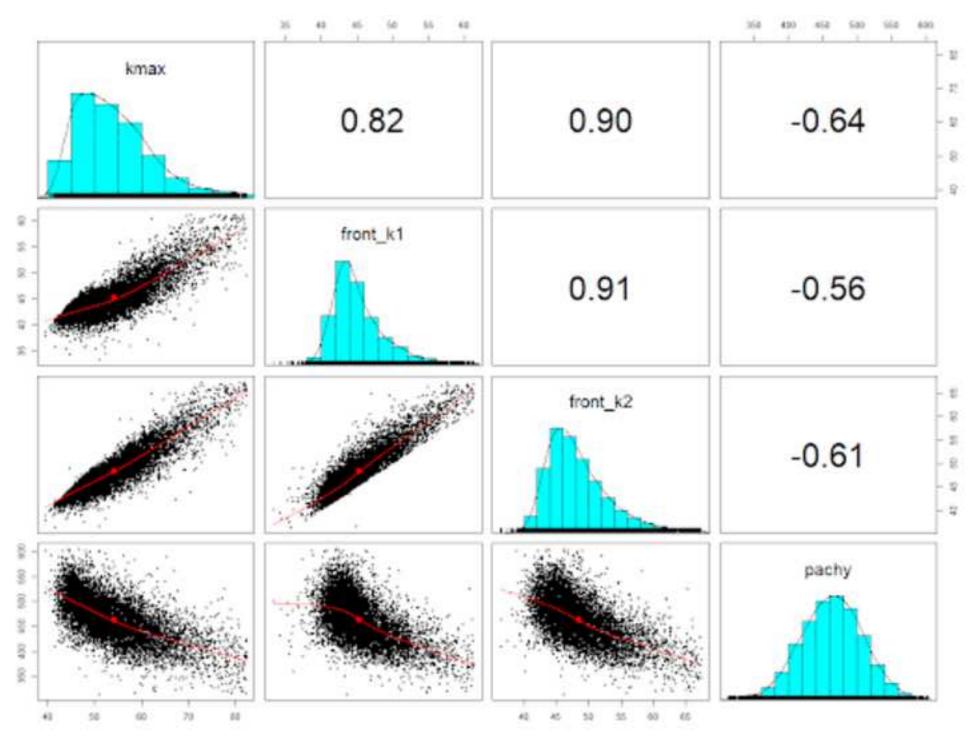
0 1 2 3 4 5 Years since first appointment Corneal crosslinking is successful in halting keratoconus progression but providing patients with a personalized visual representation of risk to progression is desirable. This research presents a model to generate projected likelihood of having crosslinking from data collected at presentation. It was trained and validated from a large dataset of 8701 eyes from keratoconus patients. Univariable and multivariable analysis was performed to identify risk factors including single nucleotide polymorphisms associated with keratoconus.

Click here to access/download;Supplementary Material: Images, Tables, Text;Supplemetary Figure 1.PNG



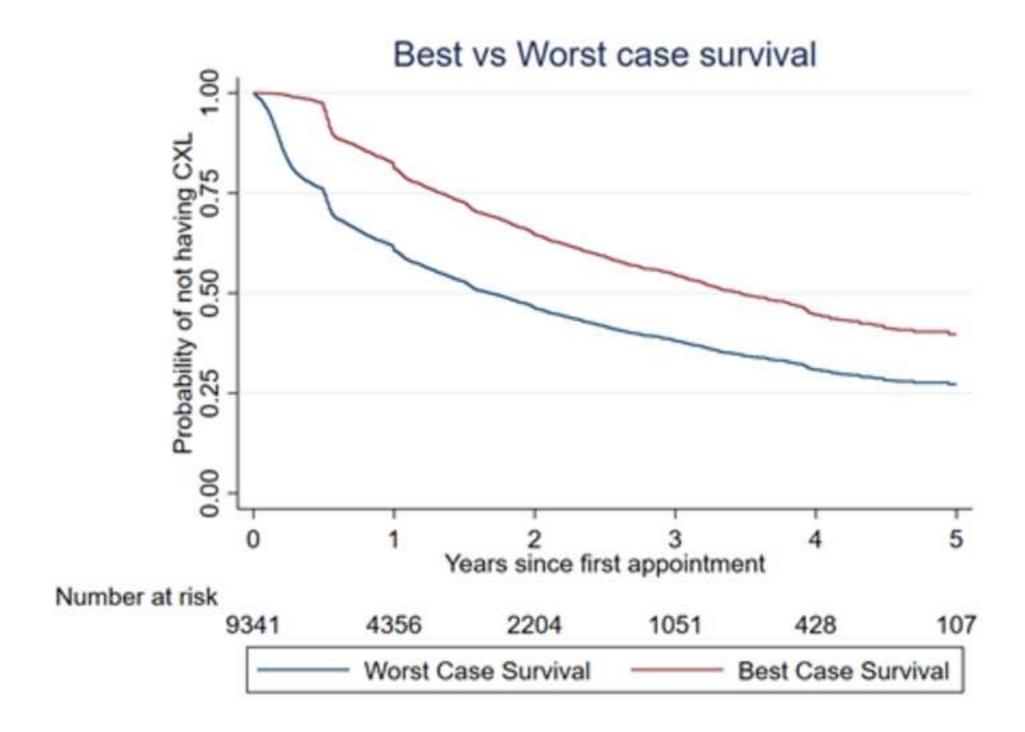
Supplementary Figure Legend 1. Geographic region grouping use in IECV. We split greater London into 6 postcode regions around its centre. Any postcodes outside of greater London were classified as the 7th region ('other'). The number of participants per region is given in Table G.

Click here to access/download;Supplementary Material: Images, Tables, Text;Supplementary Figure 2.png



Supplementary Figure 2: Correlations amongst anterior keratometry covariates and pachymetry. As explained in the results section, Front K1 and Front K2 are very highly correlated. This correlation is exemplified by the fact that removing one or other of these covariates has negligible effect on the explained variation of the model.

Abbreviations: kmax, maximum Keratometry; front_k1, flattest anterior keratometry; front_k2, steepest anterior keratometry; pachy, thinnest point pachymetry



Supplementary Figure 3: Kaplan Meier time-to-event curves for keratometric progression in the best (censored at CXL date) and worst (progressed at CXL date) case scenarios.

Supplementary Table 1. Defining disease progression in early and moderate/advanced keratoconus prior to corneal cross-linking.

Early (Kmax < 55D)	Moderate/advanced (Kmax ≥ 55D)
(1 or more)	(1 or more)
 ≥ 1 D increase Kmax ≥ 1 D increase front K1 or K2 ≥ 0.5 D increase back K2 ≥ 16 µm decrease minimum corneal thickness 	 ≥ 2.5 D increase Kmax ≥ 2.5 D increase front K1 or K2 ≥ 22 μm decrease minimum corneal thickness

Abbreviation: Kmax, maximum keratometry; K1, flat keratometry in the central 3 mm zone; K2, steep keratometry in central 3 mm zone;

Supplementary Table 2: Choice of scale and degrees of freedom (DF) using categorized (n=5) continuous covariates. Here it can be seen that 5 DF and the proportional odds model has the optimal (lowest) values for both AIC and BIC. However, we chose to model on the proportional hazards scale because our aim was to build a model that clinicians are familiar with.

	Hazard	Hazard Odds		Normal		
DF	AIC	BIC	AIC	BIC	AIC	BIC
1	953.51	1113.8	724.22	884.51	683.78	844.06
2	483.05	649.06	416.68	582.69	511.2	677.21
3	330.09	501.82	307.59	479.32	512.5	684.24
4	305.77	483.23	234.19	411.65	283.45	460.91
5	67.06	250.25	25.8	208.98	88.06	271.24
6	108.82	297.72	62.16	251.07	119.89	308.8

Abbreviations: DF, Degrees of freedom; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.

Supplementary Table 3: Multiple chained imputation results.

Covariate	Hazard Ratio [CI 95%]	P Value
Age at presentation	0.91 [0.90; 0.92]	<0.001
Kmax	1.10 [1.09; 1.11]	<0.001
Pachymetry 10 ^a	0.97 [0.96; 0.98]	<0.001
Front K1	0.95 [0.93; 0.98]	<0.001
Front K2	0.94 [0.91; 0.96]	<0.001

Abbreviations: CI, confidence interval; Kmax, maximum Keratometry; Front K1, flattest anterior keratometry; Front K2, steepest

anterior keratometry

³Minimum pachymetry in steps of 10μm.

Supplementary Table 4: Univariable results for genetic data. This includes the 28 most significantly associated Single-nucleotide Polymorphisms (SNPs)¹.

SNP name	Hazard Ratio [95% CI]	P Value	R ² _D	D
rs59771807	1.13 [0.72; 1.76]	0.60	0.00	0.08
rs3115850	1.11 [0.96; 1.27]	0.16	0.00	0.12
rs114525117	1.1 [0.85; 1.43]	0.46	0.00	0.08
rs950122	0.97 [0.85; 1.12]	0.71	0.00	0.03
rs115991721	0.72 [0.45; 1.14]	0.16	0.01	0.24
rs55678698	0.93 [0.65; 1.35]	0.72	0.00	0.06
rs6657440	1.03 [0.91; 1.17]	0.63	0.00	0.04
rs72631889	0.4 [0.22; 0.72]	0.002	0.09	0.66
rs12184325	1.09 [0.83; 1.42]	0.53	0.00	0.07
rs74460547	1.2 [0.6; 2.4]	0.60	0.00	0.12
rs72631887	1.29 [0.8; 2.07]	0.29	0.01	0.18
rs4040617	1.11 [0.96; 1.28]	0.17	0.00	0.11
rs116390263	0.96 [0.62; 1.49]	0.85	0.00	0.03
rs116452738	0.66 [0.24; 1.86]	0.43	0.01	0.17
rs3131972	1.06 [0.93; 1.21]	0.35	0.00	0.08
rs3131962	1.08 [0.94; 1.24]	0.29	0.00	0.09
rs13303222	1.02 [0.86; 1.2]	0.85	0.00	0.01

rs4970382	1.04 [0.93; 1.18]	0.48	0.00	0.06
rs116587930	1.26 [0.91; 1.75]	0.17	0.01	0.18
rs11240779	1.11 [0.98; 1.26]	0.10	0.00	0.14
rs4970459	0.98 [0.84; 1.14]	0.79	0.00	0.03
rs12562034	1 [0.84; 1.2]	0.98	0.00	0.00
rs116720794	1.09 [0.84; 1.42]	0.52	0.00	0.07
rs13303101	1.1 [0.94; 1.29]	0.24	0.00	0.11
rs4970383	1.06 [0.93; 1.2]	0.37	0.00	0.08
rs79373928	1.45 [0.82; 2.58]	0.21	0.01	0.24
rs192998324	1.15 [0.72; 1.83]	0.57	0.00	0.10
rs57181708	1.05 [0.88; 1.26]	0.55	0.00	0.04

Abbreviations: N, number of eyes; R², explained variation; D, Royston and Sauerbrei's D statistic (used as a measure of discrimination); CI, confidence interval.

Supplementary Table 5. Hazard ratios for genetic sub-analysis model using 1144 eyes containing clinical data and rs72631889 SNP.

Covariate	Hazard Ratio [95% CI]	P Value
Age at presentation	0.92 [0.90; 0.93]	<0.001
Kmax	1.10 [1.07; 1.12]	<0.001
Back K2	1.60 [1.36; 1.87]	<0.001
Pachymetry10 ^a	0.96 [0.94; 0.99]	0.010
rs72631889	0.42 [0.23; 0.77]	0.005

Abbreviations: CI, confidence interval; Back K2, steepest posterior keratometry; Kmax, maximum Keratometry; Back K2, steepest posterior keratometry

Minimum pachymetry in steps of 10µm.

Supplementary Table 6. The effect on explained variation and discrimination of removing individual covariates from the full model. Removing age has the most dramatic effect, reducing $R_{2_{D}}$ from 33.3% to 19.7%. Removing Kmax reduced $R_{2_{D}}$ from 33.3% to 28.3%. The other covariates are less important with pachymetry, Front K1 and Front K2 reducing $R_{2_{D}}$ by <1% when removed.

Covariate Removed	R^{2} D	D
None	0.333	1.45
Age at presentation	0.197	1.01
Kmax	0.284	1.29
Front K1	0.324	1.42
Front K2	0.327	1.43
Pachymetry10 ^ª	0.324	1.42

Abbreviations: R², explained variation; D, Royston and Sauerbrei's D statistic (used as a measure of discrimination); Kmax, maximum Keratometry; Front K1, flattest anterior keratometry; Front K2, steepest anterior keratometry -Minimum pachymetry in steps of 10µm.

ŧ

Region	Ν	CXL events	Dĸ	D _(k)	D k - D (k)	D S.E.
1	2,088	727	1.58	1.38	0.19	0.09
2	946	330	1.31	1.44	-0.12	0.15
3	1,014	366	1.35	1.44	-0.09	0.11
4	711	222	1.59	1.41	0.17	0.16
5	1,050	203	1.60	1.40	0.21	0.14
6	574	216	1.28	1.44	-0.16	0.14
7	2,318	1,168	1.22	1.48	-0.26	0.08

Supplementary Table 7. Internal-external cross-validation across different geographic regions.

Abbreviations: N, number in region k; $D_{\scriptscriptstyle (k)}$, discrimination omitting region k; $D_{\scriptscriptstyle k}$, discrimination predicted in region k; $D_{\scriptscriptstyle k}$ - $D_{\scriptscriptstyle (k)}$, difference in discrimination between; S.E., Standard error for $D_{\scriptscriptstyle k}$ - $D_{\scriptscriptstyle (k)}$

<u>±</u>

Covariate	Hazard Ratio	P Value	95% confidence interval	
age	0.95	<0.001	0.95	0.96
kmax	1.08	<0.001	1.06	1.10
k2	0.91	<0.001	0.88	0.93

Supplementary Table 8. Best Case Hazard Ratios

Covariate	Hazard Ratio	P Value	95% confidence interval	
age	0.93	<0.001	0.92	0.94
kmax	1.09	<0.001	1.08	1.10
k1	0.98	0.02	0.96	1.00
k2	0.95	<0.002	0.93	0.97
pachymetry	0.97	<0.003	0.96	0.98

Supplementary Table 9. Worst Case Hazard Ratios

Supplementary Text 1. Choice of Scale and Degrees of Freedom

Royston-Parmar models can be fit on a number of different scales (Hazard, Odds, Normal) and the degrees of freedom for the baseline spline can also take a range of integer values. Scale selection and degrees of freedom for the Royston-Parmar model was informed by inspecting the AIC, BIC as suggested by Royston et al. ¹. When selecting these two aspects of the model, it is important not to assume linear covariate effects for continuous variables so we categorized each continuous variable into 5 categories first. We subsequently iterated over the 3 different scales (Hazard, Odds, Probit) and 6 degrees of freedom to find the optimal (minimum) AIC and BIC. The results of iterating over both scale and degrees of freedom in order to guide further analysis can be seen in Supplementary Table 2.

1. Patrick Royston PL. Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model. Stata Press; 2011.

Supplementary Text 2. Multiple Imputation results

When applying multiple chained imputation, all 9,341 eyes were included in the analysis and we imputed the missing values for Front K1, Front K2, Back K1, Back K2, Kmax, pachymetry and ethnicity. We did not attempt to impute missing genetic data because the percentage missing (86%) was deemed to be too high. When we imputed missing data by repeating the model fitting process using multiple chained imputation the hazard ratios were very similar to the complete case analysis (Supplementary Table 3).