Article

A Machine-Learning Model Based on Morphogeometric Parameters for RETICS Disease Classification and GUI Development

José M. Bolarín [1], F. Cavas [2,*], J.S. Velázquez [2] and J.L. Alió [3,4]

[1] Technology Centre for IT and Communications (CENTIC), Scientific Park of Murcia, 30100 Murcia, Spain; bolarin5@hotmail.com

[2] Department of Structures, Construction and Graphical Expression, Technical University of Cartagena, 30202 Cartagena, Spain; jose.velazquez@upct.es

[3] Keratoconus Unit of Vissum Corporation Alicante, 03016 Alicante, Spain; jlalio@vissum.com

[4] Department of Ophthalmology, Miguel Hernández University of Elche, 03202 Alicante, Spain

* Correspondence: francisco.cavas@upct.es; Tel.: +34-968-328856

check for updates

Featured Application: This work presents a Graphics User Interface that applies two automated learning models based on machine-procured independent variables to assist ophthalmology professionals in keratoconus disease diagnosis and classification.

Abstract: This work pursues two objectives: defining a new concept of risk probability associated with suffering early-stage keratoconus, classifying disease severity according to the RETICS (Thematic Network for Co-Operative Research in Health) scale. It recruited 169 individuals, 62 healthy and 107 keratoconus diseased, grouped according to the RETICS classification: 44 grade I; 18 grade II; 15 grade III; 15 grade IV; 15 grade V. Different demographic, optical, pachymetric and eometrical parameters were measured. The collected data were used for training two machine-learning models: a multivariate logistic regression model for early keratoconus detection and an ordinal logistic regression model for RETICS grade assessments. The early keratoconus detection model showed very good sensitivity, specificity and area under ROC curve, with around 95% for training and 85% for validation. The variables that made the most significant contributions were gender, coma-like, central thickness, high-order aberrations and temporal thickness. The RETICS grade assessment also showed high-performance figures, albeit lower, with a global accuracy of 0.698 and a 95% confidence interval of 0.623–0.766. The most significant variables were CDVA, central thickness and temporal thickness. The developed web application allows the fast, objective and quantitative assessment of keratoconus in early diagnosis and RETICS grading terms.

Keywords: Scheimpflug; 3D cornea model; early keratoconus; Corrected Distance Visual Acuity (CDVA)

1. Introduction

Corneal tomography is a validated technology for evaluating the changes occurring in the corneal morphology of keratoconus (KC) disease [1], which allows for the control of the geometric decompensation driven by the asymmetry present while this disease progresses [2,3]. However, no agreement has been reached about the relative importance of the indices and technologies to be used to detect which patients may suffer post-surgical corneal iatrogenic ectasia when evaluating patients' suitability for refractive surgery [4–10].

Some studies have evaluated the objective efficiency of different indices for early stages of this disease in an individualized way. These studies are generally related to values of curvature, elevation and pachymetry [11], and to patterns based on the combination of indices that come from the same [12–16] or different technologies [7,9,17–20]. The evaluation of these patterns is based on subjective analysis [7,21], which is critical to gain a better discrimination capability according to clinical experience. Nevertheless, there are still significant discrepancies in terms of the relative value of the results obtained in their evaluation of the risk probability associated with disease development.

In KC, geometric decompensation occurs and causes disease progression [11], which means increased high-order optical aberrations and high irregular astigmatism values [22], whose main refractive sign is the impossible complete optical compensation of their ametropia by spherical-cylindrical lenses. Consequently, corrected visual acuity will diminish compared to individuals with no corneal pathology [23]. The scientific literature contains many classifications for the degree of KC severity [24–31]. However in clinical practice, it is difficult to handle the many indices on which these classifications are based for proper optical-geometrical evaluations of disease progression. From an optical point of view, patients show a deteriorated spectacle-corrected visual acuity during disease development, insofar as their visual performance worsens as the degree of KC severity progresses. Following this criterion, a classification of disease stages has been developed [23,32]: the so-called RETICS grading. This grading takes into account four geometrical parameters (Internal Astigmatism, RMS Coma-Like, Q_{8mm} and Pachymetry) and a functional one (Corrected Distance Visual Acuity (CDVA)) to establish five KC degrees: I to V.

This study develops and validates a Graphics User Interface (GUI) that combines two automatic learning models based on a set of independent variables with two aims: defining a new concept of risk probability associated with the development of early-stage KC and classifying disease severity according to the RETICS (Thematic Network for Co-Operative Research in Health) scale to assist ophthalmology professionals in disease management.

2. Materials and Methods

2.1. Patients

This research work was conceived as an observational comparative study. It comprised 169 eyes of 169 subjects divided into two groups. To avoid potential biases, those cases showing any other ocular comorbidity that could affect the present study parameters, who had undergone any ocular surgical procedure, or had worn contact lenses in the 4 weeks prior to the topographical evaluation, were excluded from both groups.

The first group, called the "control" group, comprised 62 healthy eyes of 62 patients (48.4% males, 51.6% females) whose ages ranged from seven to 60. The cases included in the control group were randomly selected from the refractive surgery candidates, and the data used for this study were acquired during their pre-surgical appointments, always with the same experienced technician.

The second group was formed by 107 KC candidates (63.2% males, 36.8% females) aged from 15 to 98. They were classified into five subgroups in accordance with the RETICS grading system [32].

The procedure followed for KC group diagnosis and classification was based on state-of-the-art clinical and topographical evaluations (Figure 1), including uncorrected distance visual acuity (UDVA), CDVA, manifest refraction (sphere and cylinder), slit-lamp biomicroscopy, Goldmann tonometry, fundus evaluation and ultrasonic pachymetry [33]. In all cases, pre-surgical evidence for KC was assessed: asymmetric bowtie pattern with or without skewed axes, localized stromal thickness reduction, conical protuberance at the apex, Fleischer ring, Vogt striae or anterior stromal scar.
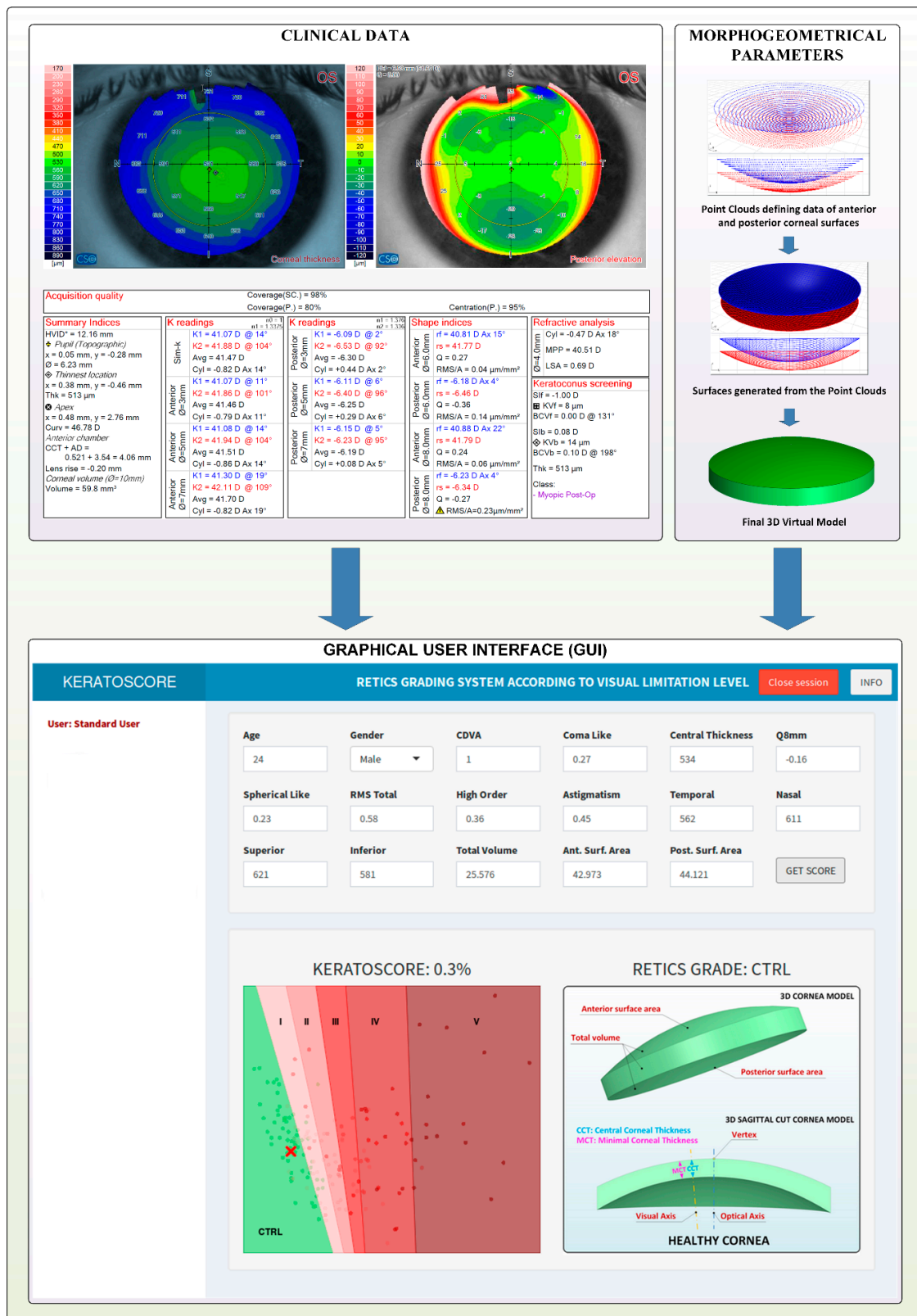
Figure 1. Methodology proposed for Keratoconus Score Calculator generation.

All the evaluations were made at Vissum Corporation Alicante (a centre affiliated with the Miguel Hernández University of Elche, Elche, Spain), and now form part of the official "Iberia" database of KC cases created for the National Network for Clinical Research in Ophthalmology RETICS-OFTARED.

Patients were adequately informed about the study and agreed to freely participate in it. The study was also ratified by the hospital's Ethical Committee for Clinical Research according to the ethical guidelines in the Declaration of Helsinki (7th revision, October 2013, Fortaleza, Brazil).

2.2. Methods

Each case selected to form part of this research was examined using a Sirius System® tomographer (Costruzione Strumenti Oftalmici, Florence, Italy), following the specifications of a validated procedure previously created by our research group, which is clearly explained in former research works [2,34]. This procedure is effective when used for both the diagnosis and characterization of KC [35,36], and comprises two phases: 3D virtual modeling, followed by a morpho-geometric analysis.

The product of this procedure is a patient-specific 3D custom corneal model, which can be studied to determine several morpho-geometric parameters that have already been described and used in a previous study [37]. Of them all, anterior corneal surface area (anterior surface area), posterior corneal surface area (posterior surface area) and total corneal volume (total volume) were selected to be used, along with demographic, pachymetry and clinical parameters (Figure 1).
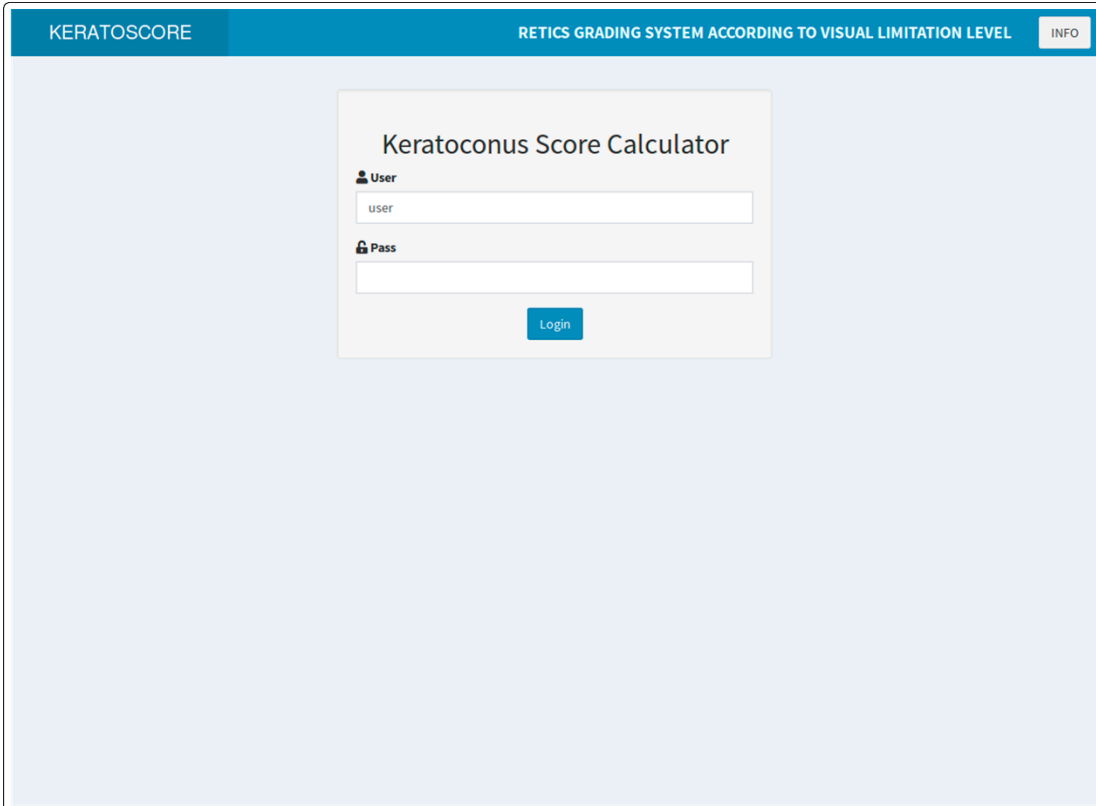
2.3. Statistical Analysis

The quantitative variables were summarized using mean±standard deviation (SD), median and interquartile range (25th and 75th percentiles). The Shapiro–Wilk test was employed to assess if the quantitative variables followed normal distribution. The Student's *t*-test was run to compare the normally distributed variables between two groups, while an ANOVA was used when there were three groups or more. For the non-normally distributed variables, the Mann–Whitney test and the Kruskal–Wallis test were, respectively, carried out. Differences in the qualitative variables among independent groups where compared by the χ^2 test. The predictive score model for early KC detection was defined using multivariate logistic regression with the control and RETICS grade I groups. Seventeen variables were included in the model: two demographic (age and gender), seven optical (CDVA, coma-like, Q_{8mm}, spherical-like, RMS total, high-order, astigmatism), five pachymetry (central thickness, temporal, nasal, superior, inferior) and three morpho-geometric (total volume, anterior surface area, posterior surface area). Model discriminative efficiency was evaluated by receiver operating characteristic (ROC) curves, when area under curve (AUC), sensitivity and specificity were taken as performance indicators. An internal cross-validation procedure was followed using bootstrap aggregating (bagging) [38]. This procedure works as follows: it first generates a new dataset of equal size by sampling with replacement from the original dataset. The model is then trained with these data. Finally, this model is used to make predictions on those cases not used during training. This procedure is repeated 100 times to obtain a set of quality parameters that can be averaged, and confidence intervals that can be calculated. On average, 63.2% of the original data were used in all these 100 training steps. The remaining 36.8% were used for validation. Ordinal multivariate logistic regression was utilized to determine the RETICS grade with the same predictor variables. In this case, the confusion matrix was employed to estimate model performance by means of sensitivity, specificity and balanced accuracy per group. Data were evaluated by R Statistics v3.6.1 (R Foundation for Statistical Computing, Vienna, Austria) [39].

Packages "tidyr", "dplyr", "dlookr" and "smbinning" were used for data loading, exploration and transformation. Packages "corrplot", "yarrr" and "FactoMiner" were employed for data visualization. Packages "pROC", "ROCR" and "Epi" were utilized for the ROC curve analysis and representation. Logistic regression models were trained with the "glm" function from the base package. Ordered logistic regression models were trained with the "polr" function from the MASS package. Confusion matrices were analyzed by the "caret" package. Optimal sample size calculations were made by the "rcompanion" and "pmsampsize" packages.

The statistical power analysis was conducted with simulation using the Wald test to estimate the power for each covariate according to sample size, as described in the literature [40,41].
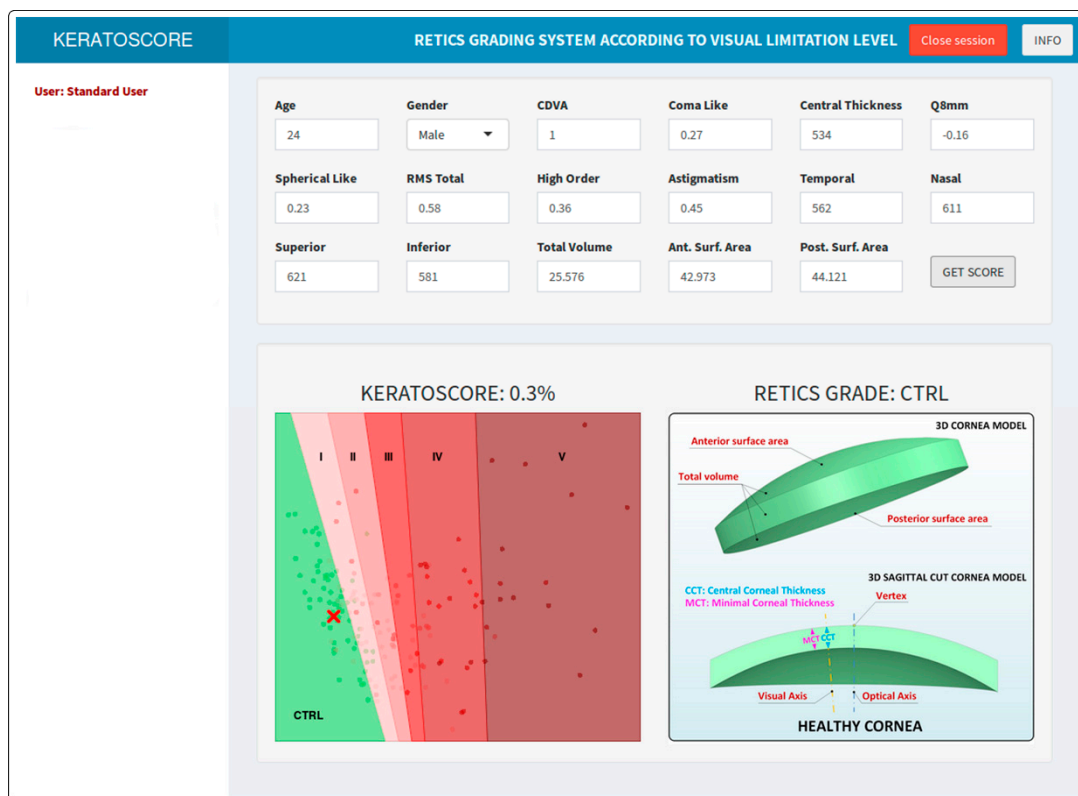
A web application was developed using Shiny v1.3.2 (RStudio Inc., Boston, MA, USA) [42] and the ShinyAuthr v0.0.99 authentication module (Paul Campbell, Paris, France) [43]. This application was deployed in a private secure institutional network (because patients' clinical data were used, and the application was, therefore, accessible only from controlled computers to minimize the security risks linked to using patients' data).

Application landing page, shown in Figure 2, is a login form that adds a secured authentication layer. No registering possibility was included, and new users can only be directly added by the administrator. After logging in, users view a form with text boxes corresponding to all the model predictors that are filled by default with sample values from a healthy individual. Users can type in new values and, after pressing the "get score" button values, they are passed to trained models and predictions are made (Figures 3–8), including the early detection KC score (known as "keratoscore"), the RETICS grade prediction and a brief graphical description of the hypothetical cornea, with a schema indicating the representation of the different parameters.
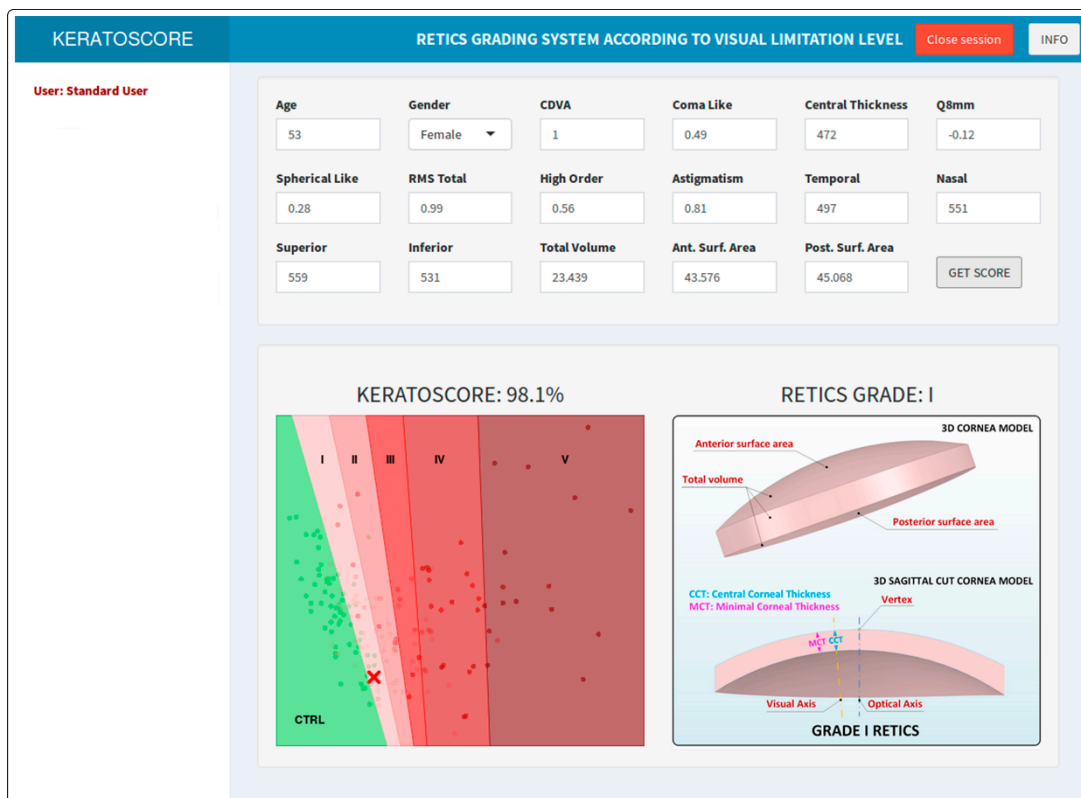

Figure 2. Application landing page showing the login form with a secured authentication layer.

Figure 3. Screenshot of healthy individuals (Ctrl) with the typical 3D virtual corneal model schematic representation.

Figure 4. Screenshot of keratoconus grade I individuals (RETICS) with the typical 3D virtual corneal model schematic representation.

Figure 5. Screenshot of keratoconus grade II individuals (RETICS) with the typical 3D virtual corneal model schematic representation.

Figure 6. Screenshot of keratoconus grade III individuals (RETICS) with the typical 3D virtual corneal model schematic representation.

Figure 7. Screenshot of keratoconus grade IV individuals (RETICS) with the typical 3D virtual corneal model schematic representation.
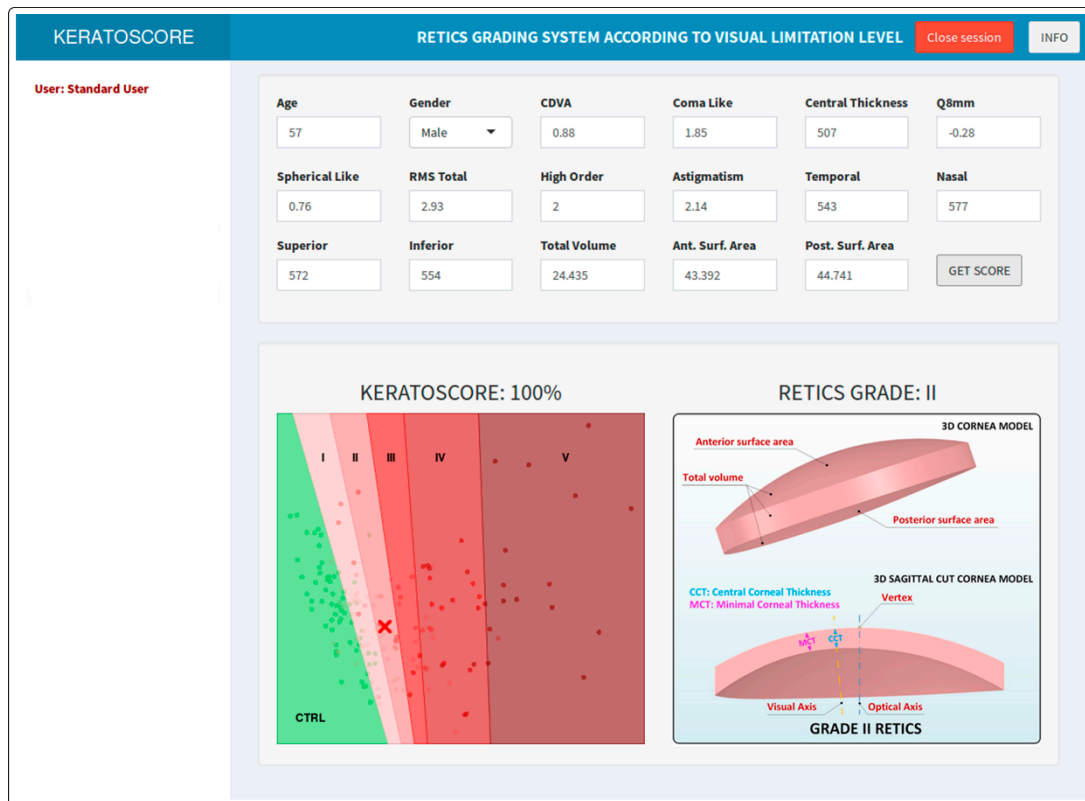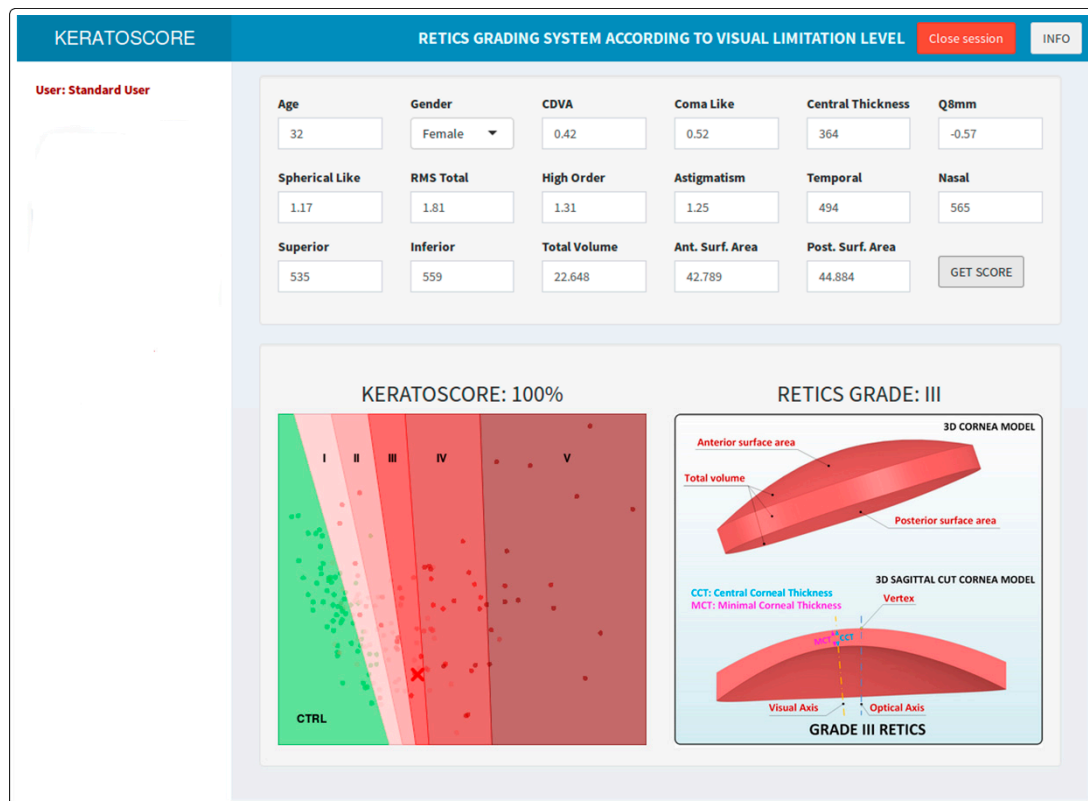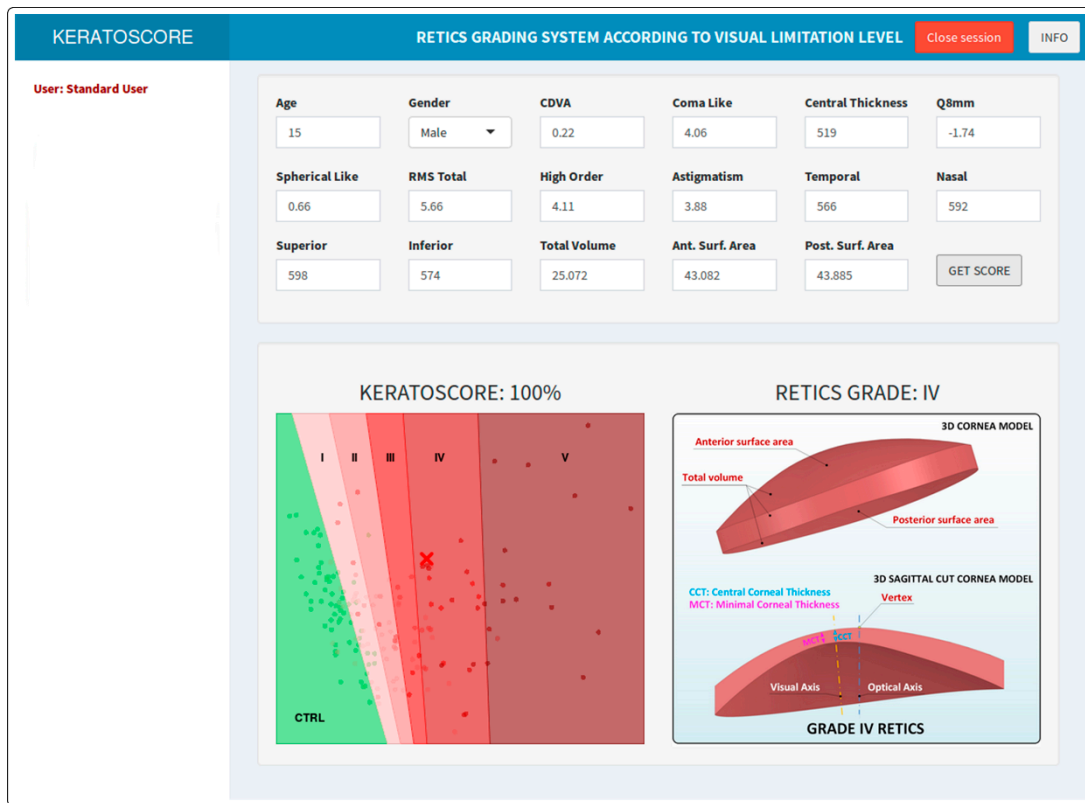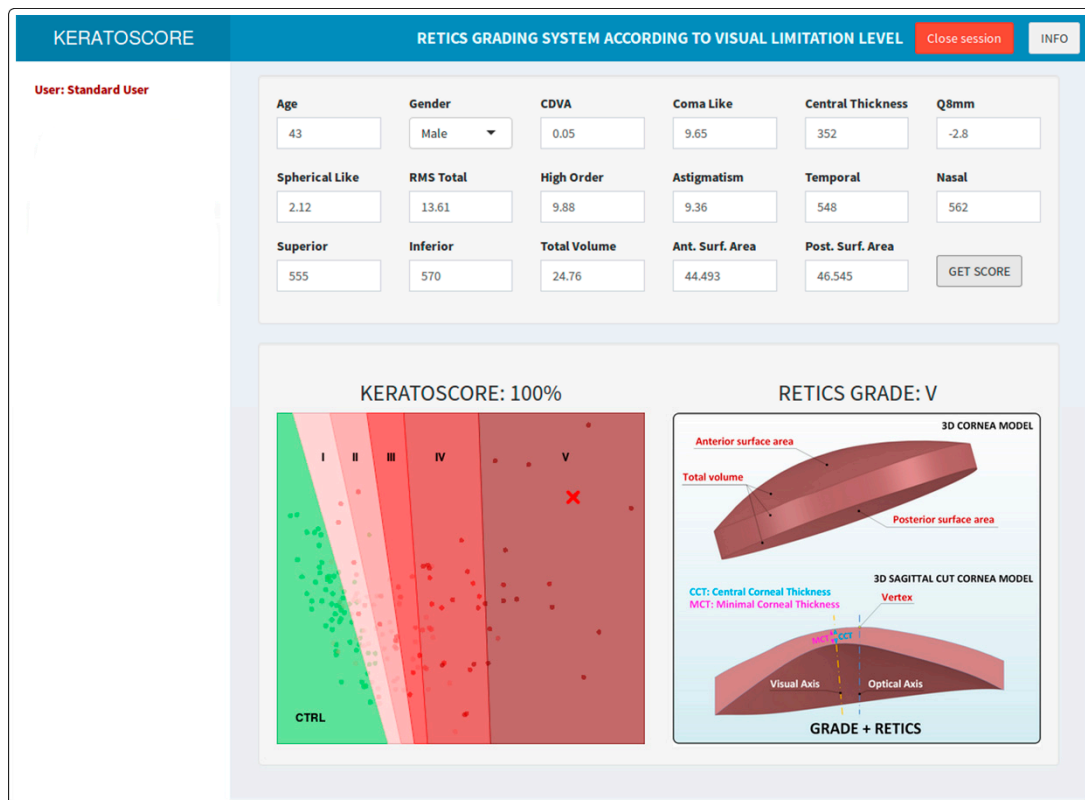

Figure 8. Screenshot of keratoconus grade V-Plus individuals (RETICS) with the typical 3D virtual corneal model schematic representation.

In summary, the GUI application was developed using a responsive design, which makes it accessible from any network-connected device because it does not require any software or driver being installed, apart from an up-to-date web browser.

3. Results

Sixty-two healthy eyes (36.7%) (control group, Ctrl) and 107 KC eyes (63.3%) (KC group, KC) were herein considered. Several subgroups were established in the KC group depending on the disease stage according to the RETICS grading system: grade I (44 eyes, 41.1%), grade II (18 eyes, 16.9%), grade III (15 eyes, 14.0%), grade IV (15 eyes, 14.0%) and grade Plus-V (15 eyes, 14.0%). The main demographic information in the control and KC subgroups are displayed and summarized in Table 1. No significant differences were found in age and gender terms among all the groups.

Table 1. Demographic information for healthy individuals (Ctrl) and the RETICS-classified keratoconus patients, graded between I and V.

	Ctrl	I	II	III	IV	V	p
Number of cases	62	44	18	15	15	15	
Age in years (mean ± SD)	37.5 ± 14.4	41.2 ± 16.6	47.4 ± 22.5	40.3 ± 12.6	34.0 ± 14.3	36.3 ± 18.0	0.161
Female/Male	32/30	12/32	6/12	7/8	6/9	4/11	0.146

Table 2 summarizes the descriptive analysis outcomes obtained for all the quantitative variable analyses in the control vs. RETICS I and the control vs. KC groups. The descriptive analysis indicated that most variables did not follow normal distribution. Statistically significant p-values were found for all the variables between the control group (Ctrl) and the RETICS grade I group, except for age ($p = 0.665$). When testing for differences within all the groups (Ctrl and RETICS grade I to V), every p-value was significant, except for age ($p = 0.344$).

Table 2. Descriptive analysis for the quantitative variables and p-values for normality and differences between RETICS groups. CDVA: Corrected Distance Visual Acuity; Q: asphericity; RMS, root mean square

Variables	Mean	SD	Median	Range	Normality	Ctrl vs. RETICS I	Ctrl–RETICS I–II–III–IV–V
Demographic							
Age	39	16	38	7–98	<0.001	0.665	0.344
Optical							
CDVA	0.77	0.32	0.96	0.05–1.20	<0.001	<0.001	<0.001
Coma-like	1.85	2.15	1.01	0.08–13.0	<0.001	<0.001	<0.001
Q8mm	−0.63	0.70	−0.45	−2.80–2.82	< 0.001	0.025	<0.001
Spherical-like	0.70	0.93	0.44	0.15–7.20	<0.001	<0.001	<0.001
RMS total	3.22	3.09	2.38	0.33–15.6	<0.001	<0.001	<0.001
High-order	2.02	2.33	1.12	0.24–13.8	<0.001	<0.001	<0.001
Astigmatism	2.31	2.25	1.57	0.04–11.22	<0.001	<0.001	<0.001
Pachymetry							
Central thickness	499	62	508	285–633	<0.001	<0.001	<0.001
Temporal	545	50	546	385–645	0.073	<0.001	<0.001
Nasal	579	48	579	451–692	0.835	<0.001	<0.001
Superior	591	50	590	408–695	0.062	<0.001	<0.001
Inferior	559	57	563	332–762	<0.001	<0.001	<0.001

Table 2. *Cont.*

Variables	Mean	SD	Median	Range	Normality	Ctrl vs. RETICS I	Ctrl–RETICS I–II–III–IV–V
Morpho-Geometric							
Total volume	24.7	1.9	24.4	19.8–29.1	0.335	<0.001	<0.001
Anterior surface area	43.35	0.59	43.24	42.49–47.44	<0.001	<0.001	<0.001
Posterior surface area	44.70	0.93	44.53	43.53–51.14	<0.001	<0.001	<0.001

The multivariate logistic regression model results are summarized in Table 3 with the coefficients for each variable. The variables that made a statistically significant contribution in the model are shown below: gender, coma-like, central thickness, high-order and temporal.

Table 3. Summary of the multivariate logistic regression model for the Ctrl vs. RETICS I patients. CDVA: Corrected Distance Visual Acuity; Q: asphericity; RMS, root mean square

		Std.	Z	p		95% CI	
Variables	Coefficient	Error	Value	Value	OR	Lower	Upper
Demographic							
Age	0.135	0.078	1.727	0.084	1.145	1.021	1.416
Gender	5.267	2.520	2.090	0.037	193.907	4.154	2.14×10^5
Clinical							
CDVA	−6.345	4.613	−1.376	0.169	0.002	0.001	1.821
Coma-like	4.072	1.917	2.124	0.034	58.687	3.615	1.16×10^4
Q8mm	6.027	4.618	1.305	0.192	414.459	0.070	3.02×10^7
Spherical-like	1.673	1.259	1.329	0.119	5.329	0.671	1.08×10^2
RMS total	2.482	1.591	1.560	0.119	11.969	1.071	8.65×10^2
High-order	5.534	2.484	2.227	0.026	253.143	5.830	3.12×10^5
Astigmatism	−1.054	1.218	−0.866	0.387	0.348	0.014	2.963
Pachymetry							
Central thickness	-0.242	0.116	−2.083	0.037	0.785	0.577	0.941
Temporal	0.028	0.012	2.248	0.025	1.028	1.009	1.064
Nasal	0.007	0.011	0.641	0.522	1.007	0.987	1.034
Superior	0.092	0.058	1.587	0.113	1.096	0.995	1.263
Inferior	−0.051	0.034	−1.479	0.139	0.951	0.870	0.995
Morpho-Geometric							
Total volume	0.067	0.547	0.124	0.901	1.070	0.349	3.908
Anterior surface area	−0.326	0.532	−0.614	0.539	0.722	0.218	2.105
Posterior surface area	−1.189	0.779	−1.525	0.127	0.305	0.044	1.199
Constant	135.027	61.595	2.192	0.028	-	-	-

The ROC curve in Figure 9 shows an optimal cut-off point of 0.475 with a training AUC of 0.990, a sensitivity of 0.977 and a specificity of 0.919, corresponding to 59 true negative cases, 41 true positive cases, three false-positive cases and three false-negative cases. The bootstrapped validation values corresponding to the 95% CI are 0.843 ± 0.058 for AUC, 0.844 ± 0.095 for sensitivity and 0.838 ± 0.081 for specificity.
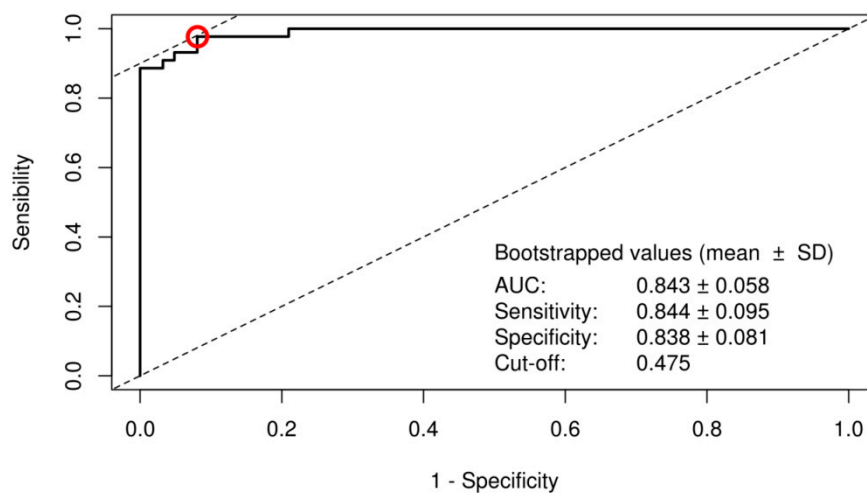
Figure 9. Logistic regression model ROC curve indicating the optimal cut-off point and the bootstrapped validation parameters.

Figure 10 shows the distribution of the calculated scores in both groups, where the vast majority of the control cases obtained a score between 0 and 0.1, and between 0.9 and 1 for RETICS I.
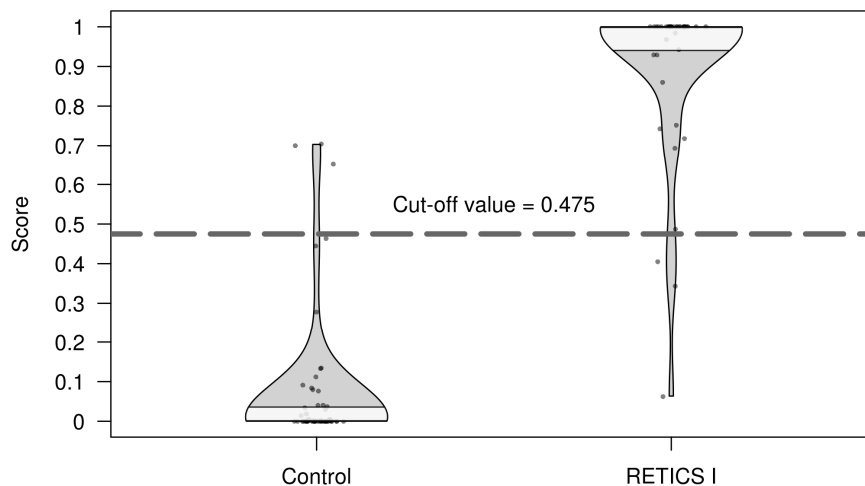
Figure 10. Distribution of the logistic regression scores between the Control and RETICS I patients, and the optimal cut-off.

The parameters for the multivariate ordinal logistic regression model are summarized in Table 4. In this case, the variables that made a statistically significant contribution to the model were: CDVA, central thickness and temporal.

The confusion matrix for the training dataset is shown in Table 5, with an overall accuracy of 0.698 at a 95% CI between 0.623 and 0.766. Sensitivity, specificity and balanced accuracy are also shown for each group.

Figure 11 depicts power according to the simulated sample size for both models, including only the variables with maximum power over 0.9 at some point within the range, plus Age and Gender. The remaining variables and their respective maxima for both the binary logistic regression and ordinal logistic regression models were Coma-Like Deviation (0.61 and 0.62), Q_{8mm} (0.55 and 0.57), Spherical-Like Deviation (0.49 and 0.55), Nasal Thickness (0.81 and 0.83), Superior Thickness (0.88 and 0.89), Inferior Thickness (0.72 and 0.77), Volume (0.69 and 0.71), Anterior Surface (0.83 and 0.89) and Posterior Surface (0.58 and 0.60). For both models, the variables of High-Order Aberration, CDVA, Central Thickness, Total RMS and Temporal Thickness have power values over 0.80 for the sample

sizes exceeding 150, and the powers for the ordinal logistic model at small sample sizes (below 100) are somewhat lower than for the binary logistic regression models.

Table 4. Summary of the ordinal logistic regression model for the Ctrl and RETICS grade I-V patients. CDVA: Corrected Distance Visual Acuity; Q: asphericity; RMS, root mean square

		Std.	*t*	*p*		95% CI	
Variables	**Coefficient**	**Error**	**Value**	**Value**	**OR**	**Lower**	**Upper**
Demographic							
Age	0.016	0.011	1.475	0.140	1.016	0.995	1.037
Gender	−0.073	0.361	−0.203	0.839	0.929	0.458	1.885
Clinical							
CDVA	−5.495	0.865	−6.355	0.001	0.004	0.001	0.022
Coma-like	0.106	0.348	0.304	0.761	1.111	0.562	2.197
Q8mm	0.184	0.359	0.512	0.609	1.201	0.595	2.423
Spherical-like	−0.063	0.250	−0.250	0.803	0.940	0.575	1.535
RMS total	−0.422	0.298	1.416	0.157	1.526	0.850	2.737
High-order	0.621	0.345	1.797	0.072	1.860	0.945	3.660
Astigmatism	−0.250	0.239	−1.046	0.296	0.779	0.488	1.244
Pachymetry							
Central thickness	−0.024	0.007	−3.298	0.001	0.977	0.963	0.991
Temporal	0.006	0.002	2.662	0.008	1.006	1.002	1.010
Nasal	−0.002	0.003	−0.771	0.441	0.998	0.992	1.004
Superior	0.006	0.007	0.933	0.351	1.006	0.993	1.019
Inferior	−0.002	0.005	−0.522	0.601	0.998	0.989	1.007
Morpho-Geometric							
Volume	0.071	0.141	0.500	0.617	1.073	0.814	1.414
Anterior area	−0.204	0.133	−1.533	0.125	0.816	0.623	1.058
Posterior area	0.046	0.129	0.361	0.718	1.048	0.814	1.348
Intercepts							
Ctrl vs. RETICS I	−16.769	0.017	−98.428	<0.001	-	-	-
RETICS I vs. II	−13.926	0.396	−35.205	<0.001	-	-	-
RETICS II vs. III	−12.270	0.522	−23.530	<0.001	-	-	-
RETICS III vs. IV	−9.998	0.765	−13.072	<0.001	-	-	-
RETICS IV vs. V	−6.691	1.057	−6.330	<0.001	-	-	-

Table 5. The ordinal logistic regression confusion matrix showing sensitivity, specificity and balanced accuracy for each group. The total sum of cells by rows shows the total number of true cases present, while each column represents how many cases the model classified in that category.

|	Predicted Value					
True Value	**Ctrl**	**I**	**II**	**III**	**IV**	**V**
Ctrl	55	7	0	0	0	0
I	12	26	4	2	0	0
II	2	9	6	3	1	0
III	0	3	3	6	3	0
IV	0	1	0	1	12	1
V	0	0	0	1	1	13
Sensitivity	0.887	0.591	0.333	0.400	0.800	0.867
Specificity	0.869	0.840	0.954	0.955	0.968	0.994
Balanced accuracy	0.878	0.715	0.643	0.677	0.884	0.930
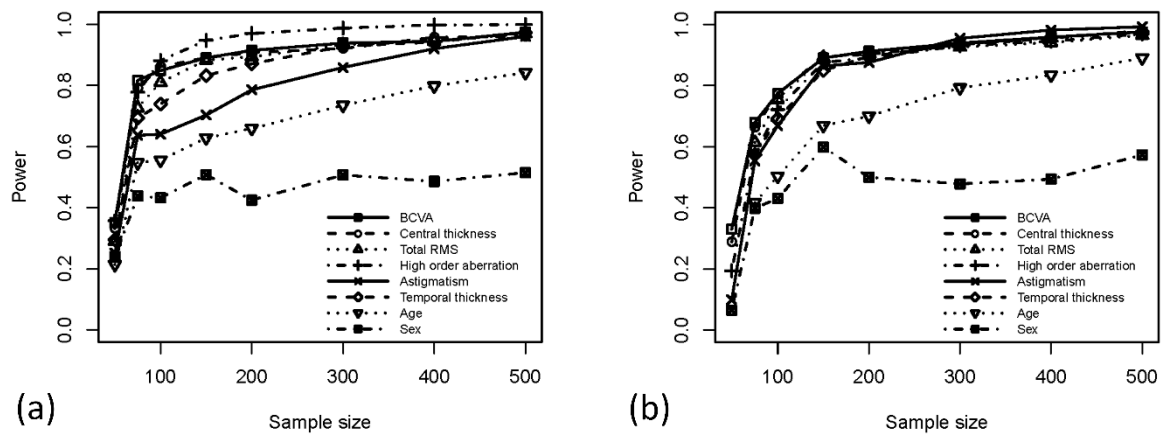
Figure 11. Power analysis results for (**a**) the binary logistic regression model; (**b**) the ordinal logistic regression model.

Graphics User Interface

Figure 3 is a screenshot corresponding to one healthy individual (Ctrl). It includes a 3D image of a characteristic cornea showing a schematic representation of how different predictors are calculated based on physical measurements.

Figures 4–8 are five screenshots corresponding to five representatives of the five RETICS grades I–V. The score rapidly rises when passing from Ctrl to grade I and remains at 100% for the other grades, which is consistent with the high sensitivity and specificity in the model. Each prediction includes a 3D image of a characteristic cornea from the corresponding RETICS group, along with a schematic representation of how different predictors are calculated based on physical measurements.

The four examples in Figure 12 indicate the difficulty of detecting early-stage KC. Indeed, some cases are wrongly classified. We must bear in mind that this early KC detection model was trained using diagnostics made by ophthalmological professionals as a "gold standard", and some inevitable undetermined amount of subjective information was taken for granted. During the fitting process, the model attempted to find some generalization to bind predictors and prediction with the best possible performance, but some samples might not match any kind of generalization given by training data's subjective nature. Therefore, it is reasonable to expect some lack of accuracy, which does not necessarily mean failure in the model's fitting capability. Our model quantitatively confirmed the difficulty of discerning both control and grade I groups as 17 of 106 cases (16%) and obtained a score between 0.1 and 0.9 which is, therefore, in the aforementioned "halfway" situation, and thus confirmed this tool's utility.
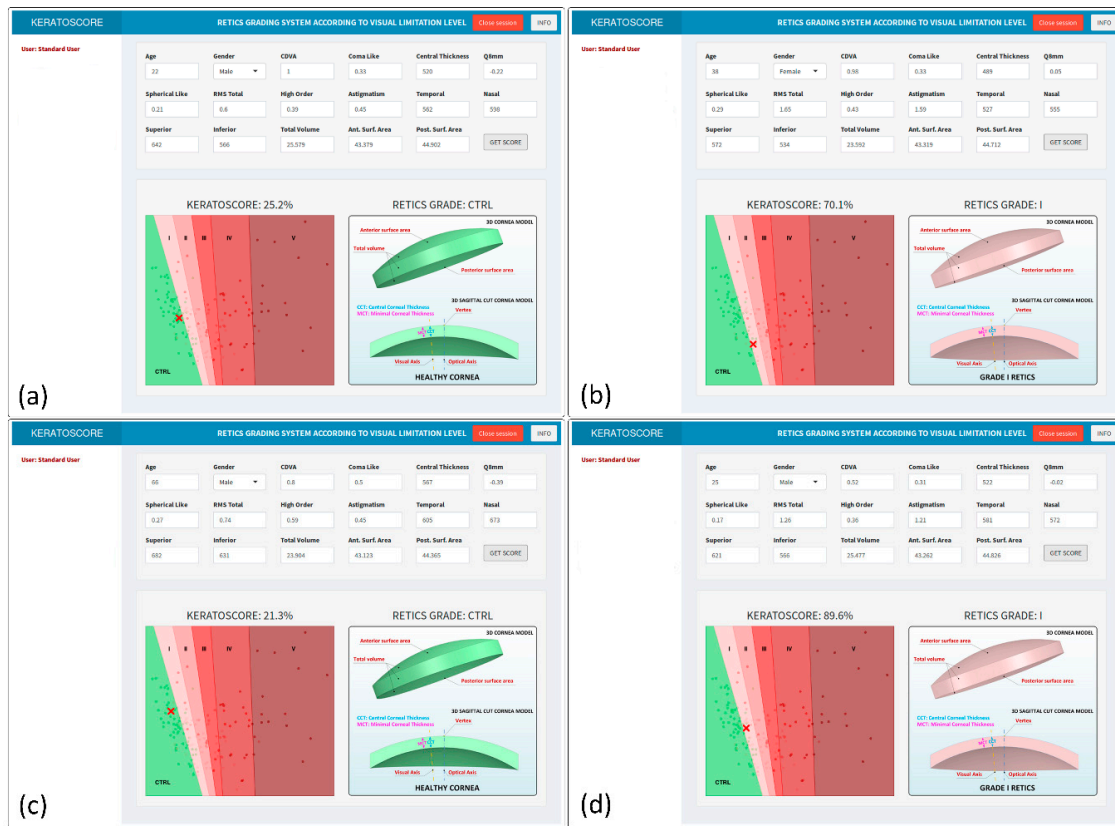
Figure 12. Screenshots showing: (**a**) a correctly classified healthy individual (true negative); (**b**) an incorrectly classified healthy individual (false-positive); (**c**) a correctly classified grade I KC individual (true positive); (**d**) an incorrectly classified grade I KC individual (false-negative).

4. Discussion

The range of techniques that allow the characterization and evaluation of the degree of KC severity vary, but the ultimate trends use machine-learning [44] and neural networks [45].

This study defines two automated learning models based on a set of independent variables (demographic, optical, pachymetric, morpho-geometric) to characterize the optical-geometrical cornea structure in different KC phases. More specifically, it considered a probability model of the risk associated with suffering early-stage KC, as well as another model to classify the degree of KC severity depending on patients' visual limitation levels.

This approach offers some very useful advantages. First, it summarizes information that derives from many parameters of different natures (qualitative or quantitative, measured in distinct units) in a single number that can be easily read and understood, which minimizes the risk of some key pieces of information going unnoticed. Otherwise, this risk is fairly high, as common analytical reports usually contain many printed pages of different parameters, which often include no associated normality range and must be read in a matter of seconds. Second, it evaluates the combined action of these different parameters, which could imply a high value in a particular key parameter, clearly indicative of disease being present, but might be less evident when the increase in many different key parameters is slight. In this situation, a score can assist health professionals in their decision-making process as it provides an objective and quantitative scale that takes into account the joint action of a set of diverse parameters [3].

A multivariate logistic regression model was fitted using healthy (Ctrl) and RETICS grade I individuals, which always made predictions ranging between 0 and 1 when applied to new data. This prediction came closer to 0 when input data were similar to those that characterized healthy controls, and close to 1 if they were similar to grade I KC patients. All those individuals with an intermediate

score were considered to be more or less halfway between both groups, i.e., they showed some kind of abnormality vs. healthy controls, which indicates KC-developing patients, but had not yet reached the cut-off point to be diagnosed as diseased. Therefore, this multivariate logistic regression output could be considered to be proportionally associated with the probability of suffering KC. A higher model output indicated a stronger probability. Even though this relationship is not linear, such scores are frequently used as predicting tools to assess patient status and prognosis in different fields [3,46–48].

For initial KC development, the analyzed model-based studies generally showed a good discrimination between normal and early KC eyes. However, one of the main challenges faced by ophthalmologists today is that no consensus has been reached by experts about the characterization patterns of early corneal ectasia [7,49]. This is mainly the result of the diversity of designations used to refer to subtler KC manifestations. It is also due to the fact that the wide variety of indices used for detection are technology-specific, which does not make them easily interchangeable [19,49]. As it is difficult to directly compare these studies, we resorted to using authors' own writings to compare them to our work.

Several studies have reported similar results to ours about models based on Scheimpflug metrics [5–7,20]. Hwang et al. [7] proposed a first model based on the combination of five metrics (AUC: 0.86, Sensitivity: 83%, Specificity: 83%). Similar results have been obtained by other authors [5,6,20] in the model development phase and based on the exclusive metrics of the same Scheimpflug technology. In contrast, our study obtained higher performance values for model development (AUC: 0.95) and similar ones in the validation phase (AUC: 0.85), and was composed of metrics that are not exclusive to the same Scheimpflug technology.

Smadja et al. [50] used a machine-learning algorithm based on decision trees to analyze 55 parameters deriving from anterior and posterior corneal measurements. They found that the most discriminant variables related to posterior surface asymmetry and thickness spatial distribution achieved 93.6% sensitivity and 97.2% specificity when discriminating between normal and forme fruste KC. However, some authors [51] considered that including eyes from patients who had already been diagnosed with KC in one eye was inappropriate because it biases the sample if we contemplate that the genetic determinants for KC appearing are already present in them.

In our study, the performance measurements of the early detection model in AUC, specificity and sensitivity terms indicated very high performance with the training dataset, with all three values reaching around 0.95, and very few false positives and false negatives. These figures are significantly lower for the validation procedure, which dropped to about 0.85, but were well over 80% in all cases. These findings indicate good model-validated performance, but also suggest the presence of some overfitting. This is otherwise reasonable if we take into account the relatively few training cases (62 healthy individuals and 44 RETICS grade I patients), which were significantly lower than the training values.

Other authors have proposed using a multivariate system based on combining two different technologies. Saad et al. [9] combined two technologies to propose a model based on 54 variables and six discriminant functions, and reported 93% sensitivity and 92% specificity in the model development phase. It was validated in a later study with 92% sensitivity and 96% specificity [52]. Other studies have suggested combining several different technologies [53–55], but the authors defined a more advanced form of KC when they included patients with manifest inferior steepening [49,56].

The only reference found in the scientific literature with a classification system that uses visual acuity as a parameter is that by Wisse et al. [3], who established a scoring system that relied on five parameters (age, quality of vision, uncorrected distance visual acuity, refraction difference, maximum keratometry difference). However, this score did not classify disease grades, but disease progression. Consequently, its aim was to determine if crosslinking treatment would be necessary or not.

Our work presents and validates a probability model of the risk associated with suffering KC. Our research group is unaware of any previous study that combined demographic, optical, pachymetric and morpho-geometric variables successfully and in real-time to detect early KC.

Finally regarding the clinical KC development phase, experts' criteria have converged to diagnose this disease as the degree of severity of progressing KC [24]. Although there are several classifications that characterize clinical KC, to the authors' knowledge, there is no disease severity classifier based on patients' visual limitation. Our classifier is based on an ordinary logistic regression model that combines 17 variables and presents an overall accuracy of 0.698; that is, our model correctly classifies almost 70% of patients. The performance indicators for ordinal logistic regression are fairly lower, particularly for the RETICS grade I, II and III patients, with balanced accuracies ranging between 0.623 and 0.766. Grade II patients present the worst accuracy, with nine in every 18 patients being wrongly classified as grade I. Once again, a small sample size is the most probable explanation for such behavior, along with its non-homogeneity (controls and grade I patients are much more abundant that grade II–V patients). It could also indicate that these three groups are not clearly differentiated. Therefore, all the conclusions drawn from the ordinal logistic regression model must be considered very carefully. Even when the results are reasonably good, the training process should be repeated with a bigger and more homogeneous sample. It would also be desirable to validate our results with other ethnicities and populations.

Another limitation of our study is the proven dependence that clinical metrics has on the technology employed to measure it [19], which means that our results are only valid for those eyes tested with a Sirius tomographer (CSO, Florence, Italy).

5. Conclusions

A web application was developed and deployed that combines two machine-learning models to support ophthalmologic professionals: a multivariate logistic regression model for early KC predictions and an ordinal logistic regression model to assign diagnosis grades on the RETICS scale. This application has a responsive design, and it allows any sort of device to be used (computer, tablet or smartphone). It also incorporates security measurements (authentication layer and accession from intranet only). The early KC prediction model shows high-performance indicators, even though some overfitting appears, while the RETICS grading prediction model's accuracy is remarkably lower, particularly for grade I, II and III patients. In both cases, repeating the training process with a bigger sample should be considered. This falls in line with recently published recommendations for sample size calculations by multivariate prediction models [57] which, for this case, proposes an optimal sample size of 374 individuals, with a minimum of 125. An optimal figure could not be reached, given our biobank's limited database size. Moreover, even though multivariate and ordinal logistic regressions are state-of-the-art and widely used techniques for modeling biomedical research data, many other powerful artificial intelligence techniques are available (particularly deep learning techniques), and their use is strongly advised for improving the quality of results.

Author Contributions: Conceptualization, J.M.B. and F.C.; methodology, J.M.B., F.C. and J.L.A.; validation, J.S.V. and J.M.B.; analysis, F.C. and J.S.V.; research, J.M.B., F.C. and J.L.A.; resources, J.M.B. and F.C.; data curation, J.S.V. and J.L.A.; writing and preparing the original draft, J.M.B. and F.C.; writing—reviewing and editing F.C., J.S.V. and J.L.A.; supervision, J.L.A.; project administration, F.C. and J.L.A.; funding acquisition J.L.A. All authors have read and agree to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pinero, D.P. *Technologies for Anatomical and Geometric Characterization of the Corneal Structure and Anterior Segment: A Review*; Informa Healthcare: London, UK, 2015; pp. 161–170.

2. Cavas-Martínez, F.; Bataille, L.; Fernández-Pacheco, D.G.; Cañavate, F.J.F.; Alio, J.L. Keratoconus detection based on a new corneal volumetric analysis. *Sci. Rep.* **2017**, *7*, 15837. [CrossRef] [PubMed]

3. Wisse, R.P.L.; Simons, R.W.P.; van der Vossen, M.J.B.; Muijzer, M.B.; Soeters, N.; Nuijts, R.M.M.A.; Godefrooij, D.A. Clinical Evaluation and Validation of the Dutch Crosslinking for Keratoconus Score. *JAMA Ophthalmol.* **2019**, *137*, 610–616. [CrossRef] [PubMed]

4. Binder, P.S. Risk factors for ectasia after LASIK. *J. Cataract Refract. Surg.* **2008**, *34*, 2010–2011. [CrossRef]

5. Binder, P.S.; Trattler, W.B. Evaluation of a risk factor scoring system for corneal ectasia after LASIK in eyes with normal topography. *J. Refract. Surg.* **2010**, *26*, 241–250. [CrossRef]

6. Chan, C.; Ang, M.; Saad, A.; Chua, D.; Mejia, M.; Lim, L.; Gatinel, D. Validation of an Objective Scoring System for Forme Fruste Keratoconus Detection and Post-LASIK Ectasia Risk Assessment in Asian Eyes. *Cornea* **2015**, *34*, 996–1004. [CrossRef]

7. Hwang, E.S.; Perez-Straziota, C.E.; Kim, S.W.; Santhiago, M.R.; Randleman, J.B. Distinguishing Highly Asymmetric Keratoconus Eyes Using Combined Scheimpflug and Spectral-Domain OCT Analysis. *Ophthalmology* **2018**, *125*, 1862–1871. [CrossRef]

8. Randleman, J.B.; Woodward, M.; Lynn, M.J.; Stulting, R.D. Risk assessment for ectasia after corneal refractive surgery. *Ophthalmology* **2008**, *115*, 37–50. [CrossRef]

9. Saad, A.; Gatinel, D. Topographic and tomographic properties of forme fruste keratoconus corneas. *Investig. Ophthalmol. Vis. Sci.* **2010**, *51*, 5546–5555. [CrossRef]

10. Seiler, T.; Quurke, A.W. Iatrogenic keratectasia after LASIK in a case of forme fruste keratoconus. *J. Cataract Refract. Surg.* **1998**, *24*, 1007–1009. [CrossRef]

11. Cavas-Martinez, F.; De la Cruz Sanchez, E.; Nieto Martinez, J.; Fernandez Canavate, F.J.; Fernandez-Pacheco, D.G. Corneal topography in keratoconus: State of the art. *Eye Vis.* **2016**, *3*, 5. [CrossRef] [PubMed]

12. Awad, E.A.; Abou Samra, W.A.; Torky, M.A.; El-Kannishy, A.M. Objective and subjective diagnostic parameters in the fellow eye of unilateral keratoconus. *BMC Ophthalmol.* **2017**, *17*, 186. [CrossRef] [PubMed]

13. Bae, G.H.; Kim, J.R.; Kim, C.H.; Lim, D.H.; Chung, E.S.; Chung, T.Y. Corneal topographic and tomographic analysis of fellow eyes in unilateral keratoconus patients using Pentacam. *Am. J. Ophthalmol.* **2014**, *157*, 103–109. [CrossRef] [PubMed]

14. De Sanctis, U.; Loiacono, C.; Richiardi, L.; Turco, D.; Mutani, B.; Grignolo, F.M. Sensitivity and specificity of posterior corneal elevation measured by Pentacam in discriminating keratoconus/subclinical keratoconus. *Ophthalmology* **2008**, *115*, 1534–1539. [CrossRef]

15. Li, Y.; Chamberlain, W.; Tan, O.; Brass, R.; Weiss, J.L.; Huang, D. Subclinical keratoconus detection by pattern analysis of corneal and epithelial thickness maps with optical coherence tomography. *J. Cataract Refract. Surg.* **2016**, *42*, 284–295. [CrossRef] [PubMed]

16. Reddy, J.C.; Rapuano, C.J.; Cater, J.R.; Suri, K.; Nagra, P.K.; Hammersmith, K.M. Comparative evaluation of dual Scheimpflug imaging parameters in keratoconus, early keratoconus, and normal eyes. *J. Cataract Refract. Surg.* **2014**, *40*, 582–592. [CrossRef]

17. Randleman, J.B.; Lynn, M.J.; Perez-Straziota, C.E.; Weissman, H.M.; Kim, S.W. Comparison of central and peripheral corneal thickness measurements with scanning-slit, Scheimpflug and Fourier-domain ocular coherence tomography. *Br. J. Ophthalmol.* **2015**, *99*, 1176–1181. [CrossRef]

18. Reinstein, D.Z.; Archer, T.J.; Urs, R.; Gobbe, M.; RoyChoudhury, A.; Silverman, R.H. Detection of Keratoconus in Clinically and Algorithmically Topographically Normal Fellow Eyes Using Epithelial Thickness Analysis. *J. Refract. Surg.* **2015**, *31*, 736–744. [CrossRef]

19. Savini, G.; Carbonelli, M.; Sbreglia, A.; Barboni, P.; Deluigi, G.; Hoffer, K.J. Comparison of anterior segment measurements by 3 Scheimpflug tomographers and 1 Placido corneal topographer. *J. Cataract Refract. Surg.* **2011**, *37*, 1679–1685. [CrossRef]

20. Shajari, M.; Jaffary, I.; Herrmann, K.; Grunwald, C.; Steinwender, G.; Mayer, W.J.; Kohnen, T. Early tomographic changes in the eyes of patients with keratoconus. *J. Refract. Surg.* **2018**, *34*, 254–259. [CrossRef]

21. Li, X.; Rabinowitz, Y.S.; Rasheed, K.; Yang, H. Longitudinal study of the normal eyes in unilateral keratoconus patients. *Ophthalmology* **2004**, *111*, 440–446. [CrossRef]

22. Ferdi, A.C.; Nguyen, V.; Gore, D.M.; Allan, B.D.; Rozema, J.J.; Watson, S.L. Keratoconus Natural Progression: A Systematic Review and Meta-analysis of 11 529 Eyes. *Ophthalmology* **2019**, *126*, 935–945. [CrossRef] [PubMed]

23. Alio, J.L.; Pinero, D.P.; Aleson, A.; Teus, M.A.; Barraquer, R.I.; Murta, J.; Maldonado, M.J.; Castro de Luna, G.; Gutierrez, R.; Villa, C.; et al. Keratoconus-integrated characterization considering anterior corneal aberrations, internal astigmatism, and corneal biomechanics. *J. Cataract Refract. Surg.* **2011**, *37*, 552–568. [CrossRef]

24. Martinez-Abad, A.; Pinero, D.P. New perspectives on the detection and progression of keratoconus. *J. Cataract Refract. Surg.* **2017**, *43*, 1213–1227. [CrossRef]

25. Romero-Jimenez, M.; Santodomingo-Rubido, J.; Wolffsohn, J.S. Keratoconus: A review. *Cont. Lens Anterior Eye* **2010**, *33*, 157–166. [CrossRef]

26. Belin, M.W.; Duncan, J.K. Keratoconus: The ABCD Grading System. *Klin. Mon. Augenheilkd.* **2016**, *233*, 701–707. [CrossRef] [PubMed]

27. Alio, J.L.; Shabayek, M.H. Corneal higher order aberrations: A method to grade keratoconus. *J. Refract. Surg.* **2006**, *22*, 539–545. [CrossRef] [PubMed]

28. Smadja, D. Topographic and tomographic indices for detecting keratoconus and subclinical keratoconus: A systematic review. *Int. J. Keratoconus Ect. Corneal Dis.* **2013**, *2*, 60. [CrossRef]

29. McMahon, T.T.; Szczotka-Flynn, L.; Barr, J.T.; Anderson, R.J.; Slaughter, M.E.; Lass, J.H.; Iyengar, S.K. A new method for grading the severity of keratoconus: The Keratoconus Severity Score (KSS). *Cornea* **2006**, *25*, 794–800. [CrossRef] [PubMed]

30. Kanellopoulos, A.J.; Asimellis, G. Revisiting keratoconus diagnosis and progression classification based on evaluation of corneal asymmetry indices, derived from scheimpflug imaging in keratoconic and suspect cases. *Clin. Ophthalmol.* **2013**, *7*, 1539–1548. [CrossRef]

31. Krumeich, J.H.; Daniel, J.; Knülle, A. Live-epikeratophakia for keratoconus. *J. Cataract Refract. Surg.* **1998**, *24*, 456–463. [CrossRef]

32. Vega-Estrada, A.; Alio, J.L.; Brenner, L.F.; Javaloy, J.; Plaza Puche, A.B.; Barraquer, R.I.; Teus, M.A.; Murta, J.; Henriques, J.; Uceda-Montanes, A. Outcome analysis of intracorneal ring segments for the treatment of keratoconus based on visual, refractive, and aberrometric impairment. *Am. J. Ophthalmol.* **2013**, *155*, 575–584.e571. [CrossRef] [PubMed]

33. Huseynli, S.; Salgado-Borges, J.; Alio, J.L. Comparative evaluation of Scheimpflug tomography parameters between thin non-keratoconic, subclinical keratoconic, and mild keratoconic corneas. *Eur. J. Ophthalmol.* **2018**, *28*, 521–534. [CrossRef] [PubMed]

34. Cavas-Martinez, F.; Fernandez-Pacheco, D.G.; De la Cruz-Sanchez, E.; Nieto Martinez, J.; Fernandez Canavate, F.J.; Vega-Estrada, A.; Plaza-Puche, A.B.; Alio, J.L. Geometrical custom modeling of human cornea in vivo and its use for the diagnosis of corneal ectasia. *PLoS ONE* **2014**, *9*, e110249. [CrossRef] [PubMed]

35. Cavas-Martínez, F.; Fernández-Pacheco, D.; Cañavate, F.; Velázquez-Blázquez, J.; Bolarín, J.; Alió, J. Study of Morpho-Geometric Variables to Improve the Diagnosis in Keratoconus with Mild Visual Limitation. *Symmetry* **2018**, *10*, 306. [CrossRef]

36. Cavas-Martinez, F.; Fernandez-Pacheco, D.G.; Cañavate, F.J.F.; Velázquez-Blázquez, J.S.; Bolarin, J.M.; Tiveron, M.; Alio, J.L. Detección del queratocono temprano mediante modelado 3D personalizado y análisis de sus parámetros geométricos. *DYNA Ing. Ind.* **2019**, *2*, 175–181.

37. Cavas-Martínez, F.; Fernández-Pacheco, D.G.; Parras, D.; Cañavate, F.J.F.; Bataille, L.; Alió, J. Study and characterization of morphogeometric parameters to assist diagnosis of keratoconus. *Biomed. Eng. Online* **2018**, *17*, 161. [CrossRef]

38. Efron, B.; Tibshirani, R. Improvements on cross-validation: The 632+ bootstrap method. *J. Am. Stat. Assoc.* **1997**, *92*, 548–560.

39. R Core Team. R: A Language and Environment for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 6 January 2020).

40. Aberson, C.L. *Applied Power Analysis for the Behavioral Sciences*, 2nd ed.; Taylor & Francis: Abingdon, UK, 2019.

41. Demidenko, E. Sample size determination for logistic regression revisited. *Stat. Med.* **2007**, *26*, 3385–3397. [CrossRef]

42. Chang, W.; Cheng, J.; Allaire, J.; Xie, Y.; Jonathan, M. Shiny: Web Application Framework for R. R Package Version 1.3.2. Available online: https://CRAN.R-project.org/package=shiny (accessed on 6 January 2020).

43. Campbell, P. Shinyauthr: Shiny Authentication Modules. R Package Version 0.0.99. Available online: https://rdrr.io/github/PaulC91/shinyauthr/ (accessed on 6 January 2020).

44. Yousefi, S.; Yousefi, E.; Takahashi, H.; Hayashi, T.; Tampo, H.; Inoda, S.; Arai, Y.; Asbell, P. Keratoconus severity identification using unsupervised machine learning. *PLoS ONE* **2018**, *13*, e0205998. [CrossRef]

45. Lavric, A.; Valentin, P. KeratoDetect: Keratoconus Detection Algorithm Using Convolutional Neural Networks. *Comput. Intell. Neurosci.* **2019**, *2019*, 9. [CrossRef]

46. Larrosa, J.M.; Moreno-Montañés, J.; Martinez-de-la-Casa, J.M.; Polo, V.; Velázquez-Villoria, Á.; Berrozpe, C.; García-Granero, M. A Diagnostic Calculator for Detecting Glaucoma on the Basis of Retinal Nerve Fiber Layer, Optic Disc, and Retinal Ganglion Cell Analysis by Optical Coherence TomographyDiagnostic Calculator of OCT for Detecting Glaucoma. *Investig. Ophthalmol. Vis. Sci.* **2015**, *56*, 6788–6795. [CrossRef]

47. Moreno-Montañés, J.; García-Nieva, A.; Osio, I.A.; Guarnieri, A.; Morilla-Grasa, A.; García-Granero, M.; Antón, A. Evaluation of RETICs Glaucoma Diagnostic Calculators in Preperimetric Glaucoma. *Transl. Vis. Sci. Technol.* **2018**, *7*, 13. [CrossRef] [PubMed]

48. Schmidl, D.; Garhöfer, G.; Schmetterer, L. A New Scoring System for Progressive Keratoconus. *JAMA Ophthalmol.* **2019**, *137*, 617. [CrossRef] [PubMed]

49. Lin, S.R.; Ladas, J.G.; Bahadur, G.G.; Al-Hashimi, S.; Pineda, R. A Review of Machine Learning Techniques for Keratoconus Detection and Refractive Surgery Screening. *Semin. Ophthalmol.* **2019**, *34*, 317–326. [CrossRef] [PubMed]

50. Smadja, D.; Touboul, D.; Cohen, A.; Doveh, E.; Santhiago, M.R.; Mello, G.R.; Krueger, R.R.; Colin, J. Detection of subclinical keratoconus using an automated decision tree classification. *Am. J. Ophthalmol.* **2013**, *156*, 237–246. [CrossRef] [PubMed]

51. Klyce, S.D. Chasing the suspect: Keratoconus. *Br. J. Ophthalmol.* **2009**, *93*, 845–847. [CrossRef] [PubMed]

52. Saad, A.; Gatinel, D. Validation of a new scoring system for the detection of early forme of keratoconus. *Age* **2012**, *37*, 37–38. [CrossRef]

53. Qin, B.; Chen, S.; Brass, R.; Li, Y.; Tang, M.; Zhang, X.; Wang, X.; Wang, Q.; Huang, D. Keratoconus diagnosis with optical coherence tomography-based pachymetric scoring system. *J. Cataract Refract. Surg.* **2013**, *39*, 1864–1871. [CrossRef]

54. Rabinowitz, Y.S.; Li, X.; Canedo, A.L.; Ambrosio, R., Jr.; Bykhovskaya, Y. Optical coherence tomography combined with videokeratography to differentiate mild keratoconus subtypes. *J. Refract. Surg.* **2014**, *30*, 80–87. [CrossRef]

55. Silverman, R.H.; Urs, R.; RoyChoudhury, A.; Archer, T.J.; Gobbe, M.; Reinstein, D.Z. Combined tomography and epithelial thickness mapping for diagnosis of keratoconus. *Eur. J. Ophthalmol.* **2017**, *27*, 129–134. [CrossRef]

56. Arbelaez, M.C.; Versaci, F.; Vestri, G.; Barboni, P.; Savini, G. Use of a support vector machine for keratoconus and subclinical keratoconus detection by topographic and tomographic data. *Ophthalmology* **2012**, *119*, 2231–2238. [CrossRef] [PubMed]

57. Riley, R.D.; Snell, K.I.; Ensor, J.; Burke, D.L.; Harrell, F.E., Jr.; Moons, K.G.; Collins, G.S. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat. Med.* **2019**, *38*, 1276–1296. [CrossRef] [PubMed]