Supplemental materials

Explainable machine learning approach as a tool to understand factors used to select the refractive surgery technique on the expert level

---

Supplementary Table 1. Questionnaire survey conducted during a preoperative evaluation. The patients determined their anticipated surgical options after fully consulted on the surgery options by an expert advisor.

Supplementary Table 2. Subjects' data variables used to construct machine learning models.

Supplementary Table 3. Detailed calculation methods of multi-categorical classification metrics including accuracy, relative classifier information (RCI), and Cohen's kappa.

Supplementary Table 4. Characteristics of the subjects in this study for training and validation data.

Supplementary Figure 1. Schematic illustration showing a multiclass one-versus-rest (OVR) classifier**.**

Supplementary Figure 2. Schematic illustration showing a multiclass one-versus-one (OVO) classifier.

Supplementary Figure 3. Schematic diagram to compare the primary factors between the explainable XGBoost model and clinician's decision.

Supplementary Figure 4. SHAP clustering force plots using the one-versus-rest XGBoost models.

Supplementary Figure 5. Examples of the features with a correlation analysis.

Supplementary Figure 6. Examples of the features with the highest importance calculated by XGBoost for two surgeons. Each machine learning model was built by one expert for each unique patient group.

Supplementary Table 1. Questionnaire survey conducted during a preoperative evaluation. The patients determined their anticipated surgical options after fully consulted on the surgery options by an expert advisor.

Name: **Patient ID:**

Age:

Gender: Male / Female

Please select items on the panel. You can select multiple items.

Order	Question	Answer 01	Answer 02	Answer 03	Answer 04	Answer 05	Answer 06
1	What was the method to correct your vision? (You can select multiple items.)	Glasses	Hard lens	Soft lens	None		
2	What kind of occupation you have? (You can select multiple items.)	Sports	Smartphone or Computer (more than 5 hours)	Driving (more than 2 hours)			
3	What kind of surgery option you anticipate? (You can select multiple items.)	LASIK	LASEK	SMILE	ICL	None	
4	What is your anticipated recovery time? (Select one item.)	One day	3 days	1 week	1 month	None	
5	What is your plan after surgery? (You can select multiple items.)	Study abroad	Employment	Military service	Other surgery	None	
6	What is major conerns about surgery? (You can select multiple items.)	Complications	Changing visual acuity	Management during recovery	Recovery duration	Budget for surgery	
7	How uncomfortable do your dry eye symptoms make you? (Select one item.)	Severe	Moderate	Mild	None		
8	Please select your past history. (You can select multiple items.)	Metabolic disease such as diabetes, hypertension, or thyroid disease	Glaucoma or Retinal disorders	Keloid or Atopic dermatitis	Recent delivery (within 3~12 months)	Other diseases	None

The questionnaire survey was originally written in Korean language and this is a translated version.

Supplementary Table 2. Subjects' data variables used to construct machine learning models.

Category	Total number	Features
Demographics & Survey	40	Age (continuous) Sex (binary) Before_Surgery_Glasses (binary) Before_Surgery_Hard_Lens (binary) Before_Surgery_Soft_Lens (binary) Before_Surgery_None (binary) Occupation_Sports (binary) Occupation_Driver (binary) Occupation_Computer_or_Smartphone (binary) Anticipated_Surgery_LASIK (binary) Anticipated_Surgery_LASEK (binary) Anticipated_Surgery_SMILE (binary) Anticipated_Surgery_ICL (binary) Anticipated_Surgery_None (binary) Anticipated_Recovery_One_Day (binary) Anticipated_Recovery_Three_Days (binary) Anticipated_Recovery_One_Week (binary) Anticipated_Recovery_One_Month (binary) Anticipated_Recovery_None (binary) Plan_After_Surgery_Study_Abroad (binary) Plan_After_Surgery_Employment (binary) Plan_After_Surgery_Military (binary) Plan_After_Surgery_Surgery (binary) Plan_After_Surgery_None (binary) Concern_Complication (binary) Concern_Visual_Acuity (binary) Concern_Management (binary) Concern_Recovery (binary) Concern_Money (binary) Concern_None (binary) Dry_Eye_Symptom_Severe (binary) Dry_Eye_Symptom_Moderate (binary) Dry_Eye_Symptom_Mild (binary) Dry_Eye_Symptom_None (binary) History_Metabolic_Disease (binary) History_Glaucoma_Or_Retinal_Disorder (binary) History_Keloid_Or_Atopic_Dermatitis (binary) History_Recent_Delivery (binary) History_Other (binary) History_None (binary)

Supplementary Table 2. Subjects' data variables used to construct machine learning models.

(continued)

Category	Total number	Features
Corneal tomography - Pentacam (both eyes)	80	Pentacam_Pupil_Diameter (continuous) Pentacam_Anterior_Chamber_Depth (continuous) Pentacam_Angle (continuous) Pentacam_Chamber_Volume (continuous) Pentacam_Keratometric_Power_Deviation (continuous) Pentacam_Corea_Volume (continuous) Pentacam_K_Max_y (continuous) Pentacam_K_max_x (continuous) Pentacam_K_max_pachy (continuous) Pentacam_Thinnest_Y (continuous) Pentacam_Thinnest_X (continuous) Pentacam_Thinnest_CCT (continuous) Pentacam_Pachy_Apex_Y_Position (continuous) Pentacam_Pachy_Apex_X_Position (continuous) Pentacam_Pachy_Apex_CCT (continuous) Pentacam_Pupil_Center_Y (continuous) Pentacam_Pupil_Center_X (continuous) Pentacam_Pupil_Center_CCT (continuous) Pentacam_Corneal_Back_Rmin (continuous) Pentacam_Corneal_Back_Rper (continuous) Pentacam_Corneal_Back_ecc (continuous) Pentacam_Corneal_Back_Astig (continuous) Pentacam_Corneal_Back_Axis (continuous) Pentacam_Corneal_Back_K_mean (continuous) Pentacam_Corneal_Back_R_mean (continuous) Pentacam_Corneal_Back_K2 (continuous) Pentacam_Corneal_Back_R_Vertical (continuous) Pentacam_Corneal_Back_K1 (continuous) Pentacam_Corneal_Back_R_Horizontal (continuous) Pentacam_Corneal_Front_Rmin (continuous) Pentacam_Corneal_Front_Rper (continuous) Pentacam_Corneal_Front_ecc (continuous) Pentacam_Corneal_Front_Astig (continuous) Pentacam_Corneal_Front_Axis (continuous) Pentacam_Corneal_Front_K_mean (continuous) Pentacam_Corneal_Front_R_mean (continuous) Pentacam_Corneal_Front_K2 (continuous) Pentacam_Corneal_Front_R_Vertical (continuous) Pentacam_Corneal_Front_K1 (continuous) Pentacam_Corneal_Front_R_Horizontal (continuous)
Ophthalmic examination (both eyes)	22	Spherical_Equivalent (continuous) Spherical_Diopter (continuous) Cylinder_Diopter (continuous) Cylinder_Axis (continuous) CDVA (logMAR) (continuous) Pupil_Diameter (continuous) IOP (continuous) CCT (continuous) Anterior_Chamber_Depth (continuous) WTW (continuous) NIBUT (continuous)
Total	142 features	

Supplementary Table 3. Detailed calculation methods of multi-categorical classification metrics including accuracy, relative classifier information (RCI), and Cohen's kappa.

Accuracy	Accuracy is a standard metric for evaluation of a classifier. It is defined as follows: $$Accuracy = \frac{\sum_i q_{ii}}{\sum_{ij} q_{ij}}$$ where the element q_{ij} refers to the number of test times and test input actually labeled C_i is C_j noted by the classifier, and these elements organize the confusion matrix. Although it is easy to notice the accuracy, it cannot give full accounts on the actual performance in multi-categorical problems.
RCI	The RCI is an entropy-based measure applicable to multi-categorical decision problems. This quantifies how much uncertainty of classification had been reduced by a machine learning classifier. It is defined as follows: $$RCI = \sum_i -\frac{\sum_j q_{ij}}{\sum_{ij} q_{ij}} log\left(\frac{\sum_j q_{ij}}{\sum_{ij} q_{ij}}\right) - \sum_j \left(\frac{\sum_i q_{ij}}{\sum_{ij} q_{ij}} \times \sum_i -\frac{q_{ij}}{\sum_i q_{ij}} log\,(\frac{q_{ij}}{\sum_i q_{ij}})\right)$$ where log refers to natural logarithm transformation. RCI represents the performance with unbalanced classes capable of distinguishing among different misclassification distributions.
Kappa	Cohen’s kappa is an alternative to classification rate that compensates for random hits. It is defined as follows: $$Kappa = \frac{\sum_{ij} q_{ij} \times \sum_i q_{ii} - \sum_{ij}(\sum_i q_{ij} \times \sum_j q_{ij})}{(\sum_{ij} q_{ij})^2 - \sum_{ij}(\sum_i q_{ij} \times \sum_j q_{ij})}$$ Kappa is a standard meter for a multi-categorical problem generally applied in several fields such as brain-computer interface.

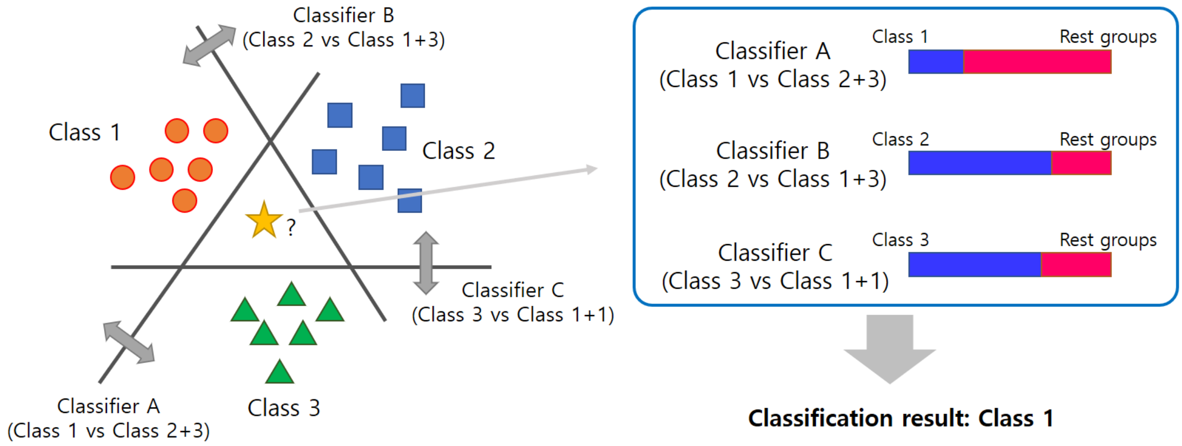
Supplementary Table 4. Characteristics of the subjects in this study for training and validation data.

Variable	**Training set (N=10,561)**	**Internal validation set (N=2,640)**	**External validation set (N=5,279)**	***P* Value[a]**
Age (years)	27.94 ± 6.12	27.89 ± 6.10	26.23 ± 6.51	<.001
Sex, female (%)	5,609 (53.1)	1,374 (52.0)	2,879 (54.5)	.081
Spherical equivalent (Diopter)	-4.56 ± 2.24	-4.55 ± 2.20	-4.80 ± 2.28	<.001
CDVA (logMAR)	-0.015 ± 0.042	-0.016 ± 0.043	0.001 ± 0.041	<.001
IOP (mmHg)	15.20 ± 4.81	15.25 ± 5.47	15.16 ± 3.06	.008
Central corneal thickness (μm)	541.86 ± 31.54	541.82 ± 31.93	542.80 ± 33.38	.070
NIBUT (s)	6.87 ± 6.60	6.90 ± 6.67	6.83 ± 5.93	<.001
Corneal refractive surgery				
LASIK	3,630 (34.4)	914 (34.6)	1,579 (29.9)	<.001
LASEK	2,891 (27.4)	729 (27.6)	1,273 (24.1)	<.001
SMILE	3,036 (28.7)	746 (28.3)	2,052 (38.8)	<.001
Contraindication cases for surgery	1,004 (9.5)	251 (9.5)	375 (7.1)	<.001

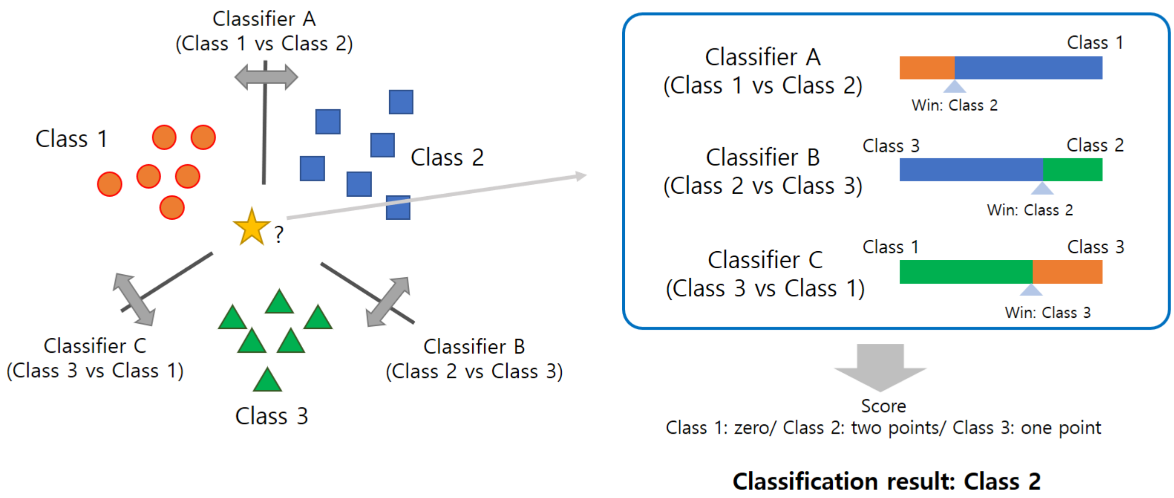
Abbreviations: CDVA, corrected distance visual acuity; IOP, intraocular pressure; LASEK, laser epithelial keratomileusis; LASIK, laser in situ keratomileusis; NIBUT, non-invasive break-up time; SMILE, small incision lenticule extraction.

[a] Comparison using the Kruskal-Wallis test and Chi-square test.

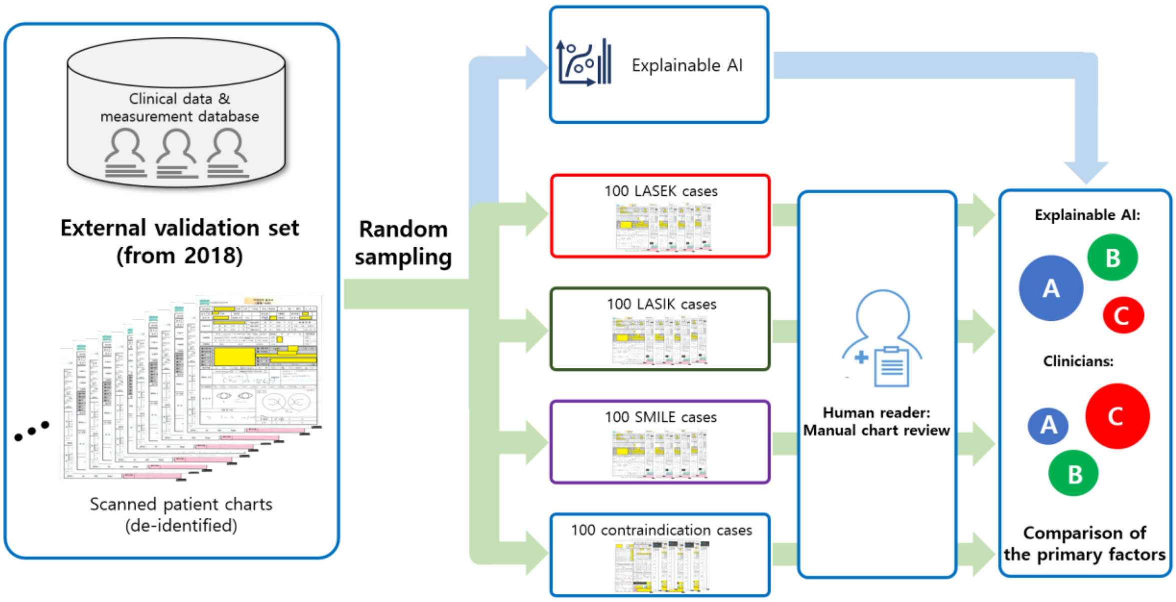
Supplementary Figure 1. Schematic illustration showing a multiclass one-versus-rest (OVR) classifier**.**



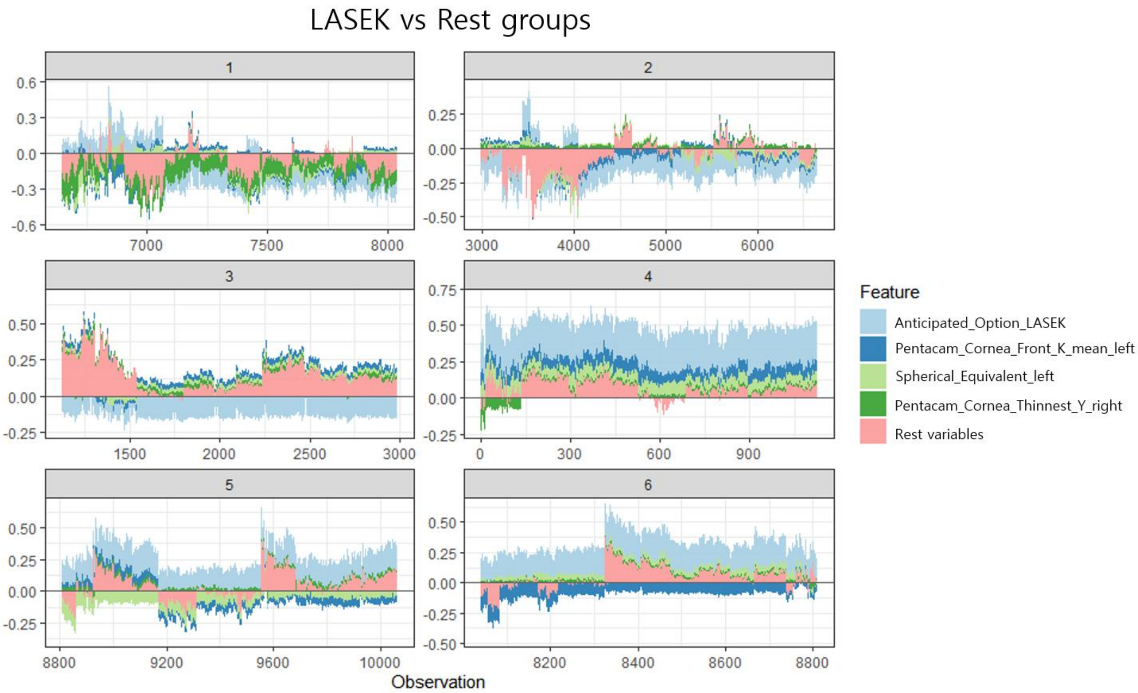
Supplementary Figure 2. Schematic illustration showing a multiclass one-versus-one (OVO) classifier.

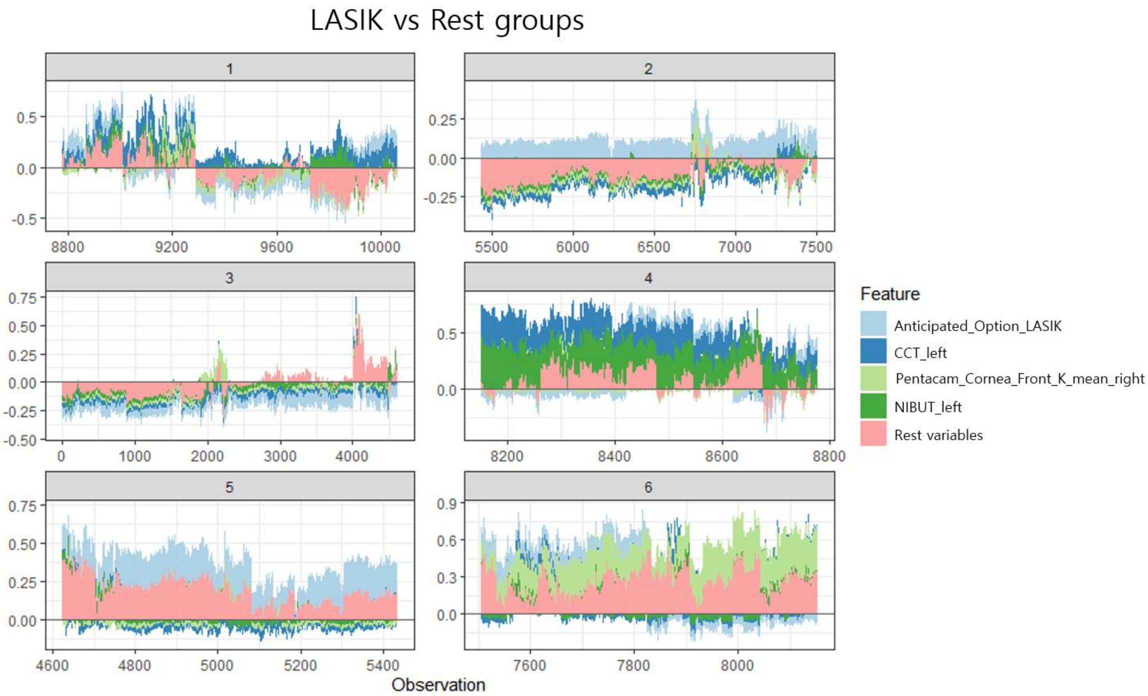


Supplementary Figure 3. Schematic diagram to compare the primary factors between the explainable XGBoost model and clinician's decision.



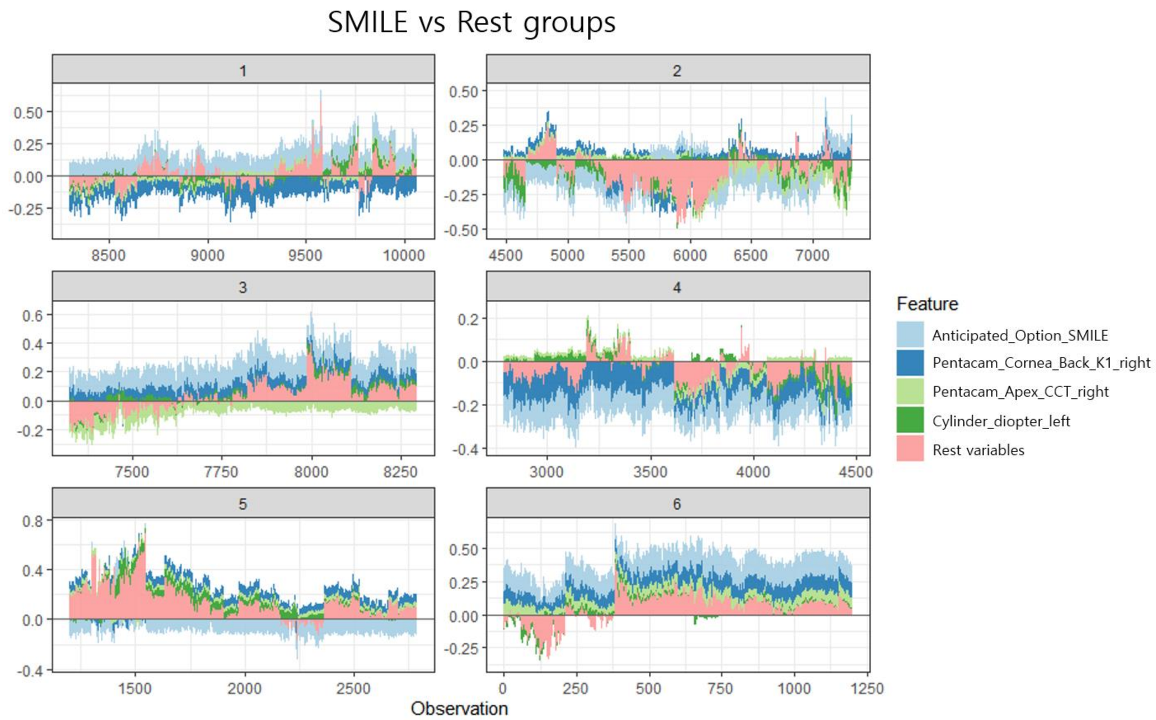
Supplementary Figure 4. SHAP clustering force plots using the one-versus-rest XGBoost models.

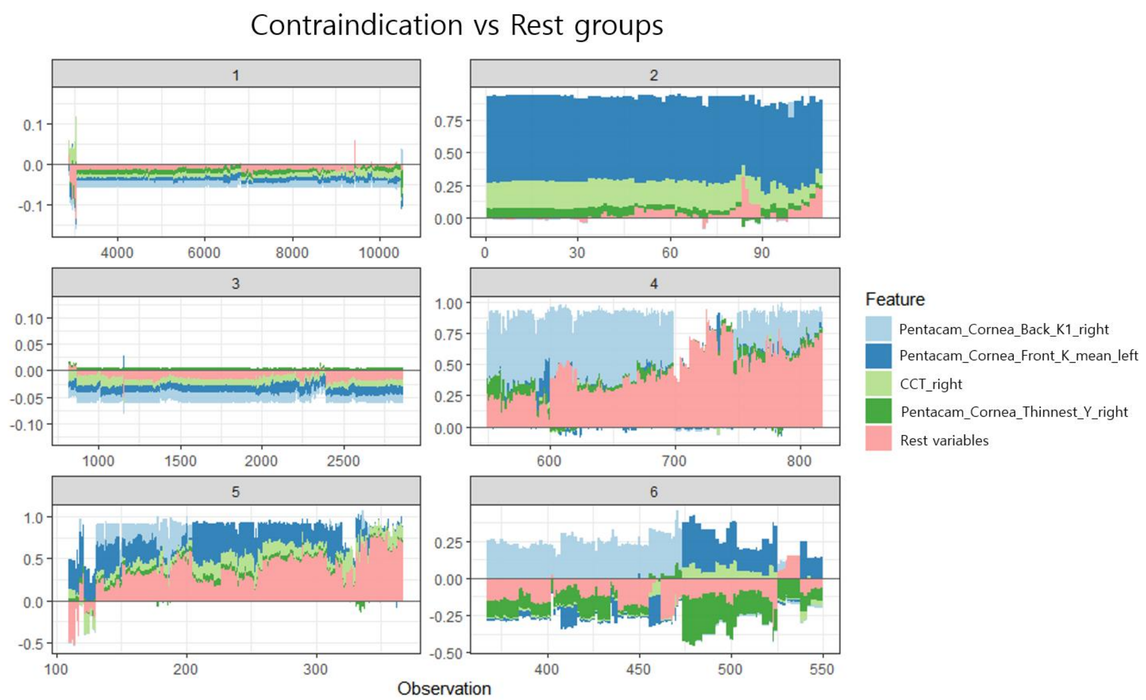




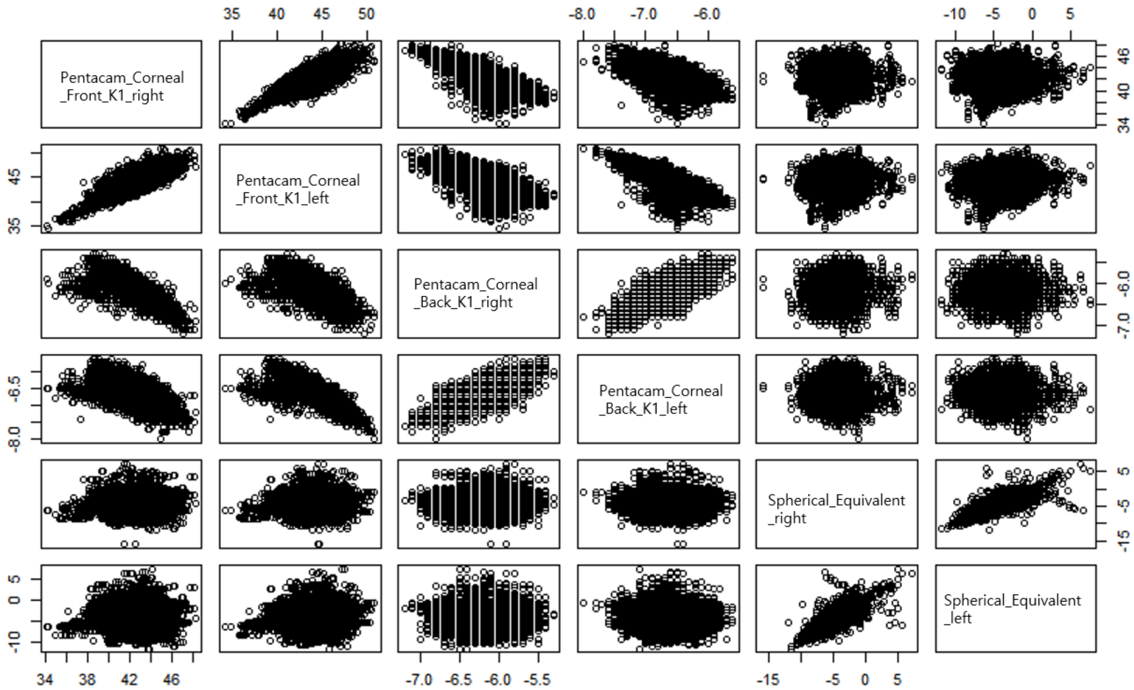
Supplementary Figure 4. SHAP clustering force plots using the one-versus-rest XGBoost models.

(continued)





Supplementary Figure 5. Examples of the features with a correlation analysis.



Supplementary Figure 6. Examples of the features with the highest importance calculated by XGBoost for two surgeons. Each machine learning model was built by one expert for each unique patient group.