A Deep Learning Approach for Successful Big-bubble Formation Prediction in Deep Anterior Lamellar Keratoplasty

Takahiko Hayashi (takamed@gmail.com)
Nihon University School of Medicine
Hiroki Masumoto
Xeno-hoc
Hitoshi Tabuchi
Hiroshima University
Naofumi Ishitobi
Tsukazaki Hospital
Mao Tanabe
Tsukazaki Hospital
Michael Grün
University of Cologne
Björn Bachmann
University of Cologne
Claus Cursiefen
University of Cologne
Sebastian Siebelmann
University of Cologne

Research Article

Keywords: deep anterior lamellar keratoplasty, area under the curve, keratoconus eyes

Posted Date: May 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-543834/v1

Abstract

The efficacy of deep learning in predicting successful big-bubble (SBB) formation during deep anterior lamellar keratoplasty (DALK) was evaluated. Medical records of patients undergoing DALK at the University of Cologne, Germany between March 2013 and July 2019 were retrospectively analyzed. Patients were divided into two groups: (1) SBB or (2) failed big-bubble (FBB). Preoperative images of anterior segment optical coherence tomography and corneal biometric values (corneal thickness, corneal curvature, and densitometry) were evaluated. A deep neural network model, Visual Geometry Group-16, was selected to test the validation data, evaluate the model, create a heat map image, and calculate the area under the curve (AUC). This pilot study included 55 patients overall (12 women, 43 men). SBBs were more common in keratoconus eyes (KC eyes) than in corneal opacifications of other etiologies (non KC eyes) (p = 0.001). The AUC was 0.746 (95% confidence interval [CI]: 0.603–0.889). The determination success rate was 78.3% (18/23 eyes) (95% CI: 56.3-92.5%) for SBB and 69.6% (16/23 eyes) (95% CI: 47.1–86.8%) for FBB. This automated system demonstrates the potential of SBB prediction in DALK. Although KC eyes had a higher SBB rate, no other specific findings were found in the corneal biometric data.

Introduction

Deep anterior lamellar keratoplasty (DALK) has been a successful treatment option for corneal stromal diseases such as keratoconus or corneal opacification [1, 2]. In comparison to the former gold standard, penetrating keratoplasty (PK), DALK has certain advantages in terms of time to visual recovery, endothelial cell density loss, and low immunological graft rejection rates [3].

However, DALK involves a steep learning curve, especially when exposing the Descemet's membrane (DM). Although different techniques such as air injection or viscoelastic devices to expose the DM have been introduced, these techniques are not always reproducible and sometimes lead to separation failure between the corneal stroma and the posterior corneal complex [4–7]. Even experienced surgeons could face a risk of failure, which could result in incomplete DM exposure or rupture. The big-bubble technique, first introduced by Dr. Anwar (Anwar's big-bubble technique), is the most popular method for exposing Descemet membranes [8]. However, despite employing this technique, a failure rate of approximately 10% for intraoperative conversion from DALK to PK due to DM rupture remains [6]. New surgical approaches, such as the microbubble incision technique and a pentacam-based approach, have reduced the conversion rate [9, 10]. Nonetheless, predicting the probability of successful big-bubble formation (SBB) before the beginning of the procedure could be extremely helpful. The success of the big-bubble technique depends on the morphology of the cornea and the extent of the pathological structure, such as the depth or extent of a scar in the corneal stroma [11]. Structural changes in the corneal tissue can be non-invasively recorded with about histological resolution using optical coherence tomography (OCT). The images thus generated are excellently suited for further image data analysis and can be collected preoperatively in almost all patients before DALK.

In recent times, machine learning analysis of clinical images, such as from OCT scans, has gained significant attention in ophthalmology and various other medical fields [12, 13]. Therefore, this study aims to evaluate the predictability of successful big-bubble formation in DALK and compare the results to preoperatively assessed biometric values.

Results

Patient Characteristics

Table 1 summarizes the patients' characteristics. In this study, 55 eyes (24 left eyes) from 55 patients were included (12 women and 43 men; mean age 46.8 ± 17.6 years [mean ± SD]). The underlying disease was keratoconus (KC eyes) in 36 patients and corneal opacification due to other etiologies (non-KC eyes) in 19 eyes. Six eyes that had been converted to PK due to intraoperative DM perforation were excluded from this study.

Table 1
Patient characteristics

	Total
Number of eyes	55
Sex: female (%)/male (%)	12 (21.8%)/43 (78.2%)
Age (years) (mean ± SD)	46.8 ± 17.6
Etiology	
Keratoconus	36
Corneal opacity	19
Grouping	
SBB (successful big bubble)	28
FBB (failed big bubble)	27

Analysis of Clinical Data

A total of 55 eyes were analyzed. Regarding the underlying disease, SBBs were more common in KC eyes than in non-KC eyes ($p = 0.001$). Although K values (Kmax or Kmean) tended to be higher in the SBB group than in the failed big-bubble (FBB) group, this difference was not statistically significant ($p = 0.101$ [SBB], $p = 0.080$ [FBB], respectively). Similarly, corneal thickness values (central corneal thickness and thinnest corneal thickness) were smaller in the SBB group, but without significant differences (p = 0.232 and p = 0.551, respectively). The SBB group tended to have a higher densitometry than the FBB group, without a significant difference (p = 0.296). A summary of biometric data is shown in Table 2.

Table 2
Comparison of preoperative biometric values between the SBB and NBB groups

	All Eyes (n = 55)	SBB Group (n = 28)	FBB Group (n = 27)	p value (FBB vs. NBB)
Etiology; proportion of KC	36/55	24/28	12/27	**0.0013†**
K max (D) (mean ± SD)	66.9 ± 15.1	71.3 ± 14.6	63.0 ± 14.5	0.101
K mean (D) (mean ± SD)	55.1 ± 11.0	58.8 ± 11.1	53.1 ± 10.2	0.080
Central corneal thickness (µm) (mean ± SD)	496 ± 186	482 ± 183	509 ± 188	0.232
Thinnest corneal thickness (µm) (mean ± SD)	412 ± 138	408 ± 129	417 ± 145	0.551
Densitometry	68.9 ± 26.3	63.1 ± 24.4	74.2 ± 26.7	0.296
SBB, successful big bubble; FBB, failed big bubble; SD, standard deviation				
†Pearson's chi-square test; *Mann–Whitney U test				

ROC Curve for Eye-by-Eye Determination

The AUC was 0.746 (95% confidence intervals [CI]: 0.603–0.889) (Fig. 1). The determination success rate was 78.3% (18/23 eyes) (95% CI: 56.3–92.5%) for SBB and 69.6% (16/23 eyes) (95% CI: 47.1–86.8%) for FBB.

Discussion

Automated image recognition using deep learning methods is becoming increasingly important in the decision making, planning, and execution of ophthalmic surgeries. While there are numerous studies on the retina, there are very few studies on the anterior eye segments, especially the cornea [14, 15]. Nevertheless, it has been shown that artificial intelligence and image data obtained by OCT can be used to detect and classify certain corneal diseases, and even predict the probability of the need for future keratoplasty [16]. Similar findings have been described before cataract surgery or for the detection of KC [17, 18]. The present study confirms that, even in complex situations such as predicting the success of a particular surgical maneuver (BB formation in DALK), intelligent algorithms can make very good predictions. KC eyes had a better success rate of BB formation than those with corneal opacification, although a selection of pre-selected, established parameters did not reveal a statistically significant difference between the SBB and FBB groups.

The limitations of this pilot study are that, owing to the small sample size, the accuracy of individual and special indications for DALK could not be elucidated, such as herpetic scars, traumatic scars, exclusively KC, or corneal dystrophies. Therefore, in the future, it would be useful to collect OCT data from several

centers to develop synergies and perform a subgroup analysis by using the here developed algorithm. This could be facilitated by cross-country and cross-device web platforms. Alternatively, large amounts of available data would overcome this limitation. In the present study, the success rate of one type of BBF (Anwar's big-bubble) by only two surgeons was evaluated. Therefore, these data are probably not easily transferable to other surgeons or techniques; nevertheless, this pilot study shows the possibilities of such analysis using artificial intelligence. However, one of the main points of criticism for the clinical implementation could be that the criteria of the algorithm used to make its relatively precise statements is not sufficiently clear. In the future, similar algorithms combined with modern intraoperative imaging technologies, such as microscope-integrated optical coherence tomography and real-time intraoperative pattern recognition algorithms could help ophthalmic surgeons not only before or after surgery, but also during the execution of surgical maneuvers [19]. We consider it possible to predict SBB using microscope images alone. Future studies should aim to integrate large amounts of data using various imaging modalities.

Materials And Methods

Study Design

In this retrospective study, the OCT images of all patients who underwent DALK between March 2013 and July 2019 were analyzed. Images were acquired using a Spectral Domain OCT device (SD-OCT, Heidelberg Engineering, Heidelberg, Germany). The study complied with the ethical standards of the Declaration of Helsinki and was approved by the institutional review board of the University of Cologne, Cologne, Germany (File Number 15–301). Informed written consent was obtained from all subjects before enrollment.

Surgical Technique

All surgeries were performed by two experienced surgeons under general anesthesia [20]. DALK was performed in a standardized fashion using Anwar's big-bubble technique as previously described [8]. Briefly, the host cornea was marked using a trephine (Katena, Denville, USA) at 7.5–8.0 mm, the DM exposed, and the host stroma carefully removed. The donor graft without DM was prepared using a donor punch (Katena, Denville, USA) at 7.75–8.5 mm and then secured with a 10–0 nylon suture by either a running suture or interrupted suture technique. In cases where no big bubbles could be created, the DM was exposed by dissecting the stroma in layers using a manual dissection technique [5].

Grouping by Intraoperative Procedures

Patients who underwent DALK were divided into two groups: (1) successful big-bubblegroup or (2) failed big-bubble group.

Collection of Preoperative Biometric Data

To obtain preoperative biometrical data, the eyes were examined preoperatively using the Scheimpflug tomography system (Pentacam HR, Oculus GmbH, Wetzlar, Germany), preoperative corneal thickness values (central corneal thickness, thinnest corneal thickness), keratometric values (Kmax, Kmin), and the corneal densitometry was thereon compared between the two groups (SBB and FBB).

Dataset and Grouping for Deep Learning

For constructing the artificial intelligence-based image analysis system, the following two categories were defined: SBB vs. FBB. The preoperative corneal cross-sectional OCT images captured from SD-OCT volume scans or manual single scans were analyzed.

In this study, the K-fold cross-validation (K = 5) method was applied [21, 22]. Image data were divided into K groups. The K-1 group was used as training data, and the remaining group as evaluation data. The process was repeated K times until all groups were used as evaluation data. During training, the following data expansion was performed for each epoch: rotation, displacement, shear, enlargement, vertical inversion, horizontal inversion, two types of brightness adjustment, two types of gamma correction, blurring, histogram flattening, and two types of noise load. Training was performed for a deep convolutional neural network.

Model Construction

For the original image, the image consisting of 1024 × 242 pixels was read as the RGB channel. First, the image was resized to 224 × 224 pixels. The pixel value consisting of 8 bits (range of 0–255) was divided by 255 and normalized to the range of 0–1.

In this study, we tested various deep neural network (DNN) models of DenseNet121, EfficientNetB0, and Visual geometry group-16 (VGG-16) and finally selected VGG-16 because of its balanced values (data not shown). This type of DNN is known to automatically learn the local features of an image to generate a classification model [23–27]. VGG-16 consists of five blocks and three fully connected layers. Each block consists of several convolutional layers that automatically extract features and max-pooling layers that reduce position sensitivity and improve generalization performance [28].

The convolution layer stride was equal to one and the layer padding was set to be the same. As a result, the convolutional layer only grasped the image features and did not downsize. The activation function was set as rectified linear unit (ReLU) to avoid the problem of gradient disappearance [29]. The stride of the max-pooling layer was set to two. After Block 5, there were flatten layers and two fully connected layers. Spatial information was removed from the feature vector extracted by the flatten layer. The fully connected layer compressed the information. Finally, the probability of each class was evaluated by passing a function called softmax, and target images classified. Fine-tuning (or full retraining) was performed to increase the learning speed so that high performance could be achieved with a small amount of data [30]. The parameters of blocks 1–4 used ImageNet as initial values.

The parameters of each layer were updated using the momentum stochastic gradient descent algorithm (updated using learning rate = 0.0005, initial term = 0.9) [31, 32]. In addition, training and verification were performed using Keras, a Python TensorFlow Application Programming Interface (API) (https://www.tensorflow.org/).

Deep Learning

Using the Score-CAM method [33], we created a heat map image showing where the DNN was concentrated. The target layer was set as the maximum pooling layer of block 5, and the ReLU function was used to correct the loss function during backpropagation.

The performance evaluation indices were the area under the curve (AUC) and the correct answer rate in the successful and unsuccessful eye. The calculation method was as follows. First, since the probability of successful DM separation can be discerned by the neural network output, in each image, an image output with a value > 0.5 was regarded as successful, whereas images with a low value were unsuccessful. When the ratio of the number of successfully determined images in each eye exceeded a certain threshold, it was determined that the eye will succeed, and the AUC was calculated by adjusting the threshold.

The ROC curve used to determine the AUC was created by defining the point at which the value used to indicate SBB positively exceeded the threshold cutoff value output from the softmax function. We created 100 ROC curves from 100 patterns with 10% thinned out; thus, this model was applied to only 90% of the test data. A total of 100 AUCs were calculated from each ROC curve, and a 95% CI was obtained by assuming a normal distribution and average standard deviation. In addition, we set the threshold value that maximized the performance using the Youden index and then calculated the correct answer rate in the successful and unsuccessful eyes [34]. The CIs of sensitivity and specificity were calculated assuming a binomial distribution.

Statistical Analysis

Statistical analyses were performed using JMP Pro software version 14.0.0 (SAS Institute, Cary, NC, USA). To compare the continuous variables in each group, we used either a one-way analysis of variance (ANOVA) or Mann–Whitney U test, while nominal variables such as patient sex and operated eye were compared using Pearson's chi-square test.

Declarations

Data Availability Statement

The data that support the findings of this study are available from the corresponding author [T.H], upon reasonable request.

Acknowledgments

None.

Author Contribution Statement

T.H.: writing, reviewing, and editing of the manuscript; H.M.: curation of the data and editing of the manuscript; H.T, and N.I.: reviewing and editing of the manuscript; M.T, and M.G.: curation and analysis of the data and investigation; B.B.: reviewing of the manuscript and investigation; C.C.: administration of the project and investigation; S.S.: supervision and validation; All authors critically checked the manuscript and approved its submission.

Competing interest

The authors declare no potential competing interest.

Funding

This study did not receive any public or private funding.

References

1. Reinhart, W. J. *et al*. Deep anterior lamellar keratoplasty as an alternative to penetrating keratoplasty a report by the American academy of ophthalmology. *Ophthalmology***118**, 209-218 (2011).
2. Flockerzi, E. *et al*. Trends in corneal transplantation from 2001 to 2016 in Germany: A report of the DOG-Section cornea and its keratoplasty registry. *J. Ophthalmol.* **188**, 91-98 (2018).
3. Hos, D., et al. Immune reactions after modern lamellar (DALK, DSAEK, DMEK) versus conventional penetrating corneal transplantation. *Retin. Eye Res*. **73**, 100768 (2019).
4. Shimazaki, J., Shimmura, S., Ishioka, M., & Tsubota, K. Randomized clinical trial of deep lamellar keratoplasty vs penetrating keratoplasty. *Am J Ophthalmol.***134**, 159-165 (2002).
5. Sugita, J., & Kondo, Deep lamellar keratoplasty with complete removal of pathological stroma for vision improvement. *Br. J. Ophthalmol*. **81**, 184-188 (1997).
6. Sarnicola, V., Toro, P., Gentile, D., & Hannush, S. B. Descemetic DALK and predescemetic DALK: outcomes in 236 cases of keratoconus. *Cornea*. **29**, 53-59 (2010).
7. Smadja, D. *et al*. Outcomes of deep anterior lamellar keratoplasty for keratoconus: learning curve and advantages of the big bubble technique. *Cornea***31**, 859-863 (2012).
8. Anwar, M., & Teichmann, K. D. Big-bubble technique to bare Descemet's membrane in anterior lamellar keratoplasty. *Cataract Refract. Surg*. **28**, 398-403 (2002).
9. Riss, S., Heindl, L. M., Bachmann, B. O., Kruse, F. E., & Cursiefen, C. Microbubble incision as a new rescue technique for big-bubble deep anterior lamellar keratoplasty with failed bubble formation. *Cornea***32**, 125-129 (2013).
10. Riss, S., Heindl, L. M., Bachmann, B. O., Kruse, F. E., & Cursiefen, C. Pentacam-based big bubble deep anterior lamellar keratoplasty in patients with keratoconus. *Cornea***31**, 627-632 (2012).

11. Ozmen, M. C. *et al*. Prediction of Descemet's membrane perforation during deep anterior lamellar keratoplasty in patients with keratoconus with stromal scar. *Eye Contact Lens***44 Suppl 2**, S176-S179 (2018).
12. Takahashi, H., Tampo, H., Arai, Y., Inoue, Y. & Kawashima, Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy. *PLoS One***12,** e0179790 (2017).
13. California Healthcare Foundation. *Diabetic Retinopathy Detection* https://www.kaggle.com/c/diabetic-retinopathy-detection
14. Hayashi, T. *et al*. A deep learning approach in rebubbling after Descemet's membrane endothelial keratoplasty. *Eye Contact Lens***46**, 121-126 (2020).
15. Maruoka, S. *et al*. Deep neural network-based method for detecting obstructive meibomian gland dysfunction with in vivo laser confocal microscopy. *Cornea***39**, 720-725 (2020).
16. S. *et al*. Predicting the likelihood of need for future keratoplasty intervention using artificial intelligence. *Ocul. Surf.***18**, 320-325 (2020).
17. Goh, J. H. L. *et al*. Artificial intelligence for cataract detection and management. *Asia Pac. J. Ophthalmol.* (Phila) **9**, 88-95 (2020).
18. Yousefı, S. *et al*. Keratoconus severity identification using unsupervised machine learning. *PLoS One* **13**, e0205998; 10.1371/journal.pone.0205998 (2018).
19. Steven, P. *et al*. Optimising deep anterior lamellar keratoplasty (DALK) using intraoperative online optical coherence tomography (iOCT). *J. Ophthalmol*. **98**, 900-904 (2014).
20. Schaub, F. *et al*. Impact of donor graft quality on deep anterior lamellar keratoplasty (DALK). *BMC Ophthalmol*. **17,** 204 (2017).
21. Mosteller, F. & Tukey, J. W. Data analysis, including statistics in *Handbook of Social Psychology: Vol. 2. Research Methods* (eds. Lindzey, G., Aronson, E.) 80-203 (Addison-Wesley in Reading, MA, 1968).
22. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proc. Int. Joint Conf. AI*. **2**, 1137-1145 (1995).
23. Simonyan, K., & Zisserman, Very deep convolutional networks for large-scale image recognition. https://arxiv.org/abs/1409.1556v4 (2014).
24. Deng, J. *et al*. Imagenet: A large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition. *Piscataway*: IEEE, 248-55 (2009).
25. Lee, C. Y., Xie. S., Gallagher, P., Zhang, Z., & Tu, Z. Deeply-supervised nets. *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics* http://proceedings.mlr.press/v38/lee15a.pdf (2015).
26. Huang, G., Liu, Z., van der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. Preprint at: arXiv:1608.06993* (2016).
27. Tan, M., & Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks Preprint at: arXiv: 1905.11946* (2019).

28. Scherer, D., Müller, A., & Behnke, Evaluation of pooling operations in convolutional architectures for object recognition. In: Artificial Neural Networks–ICANN, 2010. Berlin: Springer, 92-101.
29. Glorot, X., Bordes, A., & Bengio, Deep sparse rectifier neural networks in *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*. (eds Diamantaras, K., Duch, W., Iliadis, L. S.) 92-101 (Springer, Berlin, Heidelberg 2010).
30. Agrawal, P., Girshick, R., & Malik, Analyzing the performance of multilayer neural networks for object recognition. In: European Conference on Computer Vision, 2014. Cham: Springer, 329-344.
31. Qian, On the momentum term in gradient descent learning algorithms. *Neural Netw.***12**, 145-151 (1999).
32. Nesterov, A. method for unconstrained convex minimization problem with the rate of convergence O (1/k^2). *USSR Acad. Sci.* **269**, 543-547 (1983).
33. Wang, , Du, M., Yang, F., & Zhang, Z. Score-cam: Improved visual explanations via score-weighted class activation mapping. *arXiv Preprint at: arXiv:1910.01279*; 2019.
34. Youden, W. J. Index for rating diagnostic tests. *Cancer***3**, 32-35 (1950).

Figures

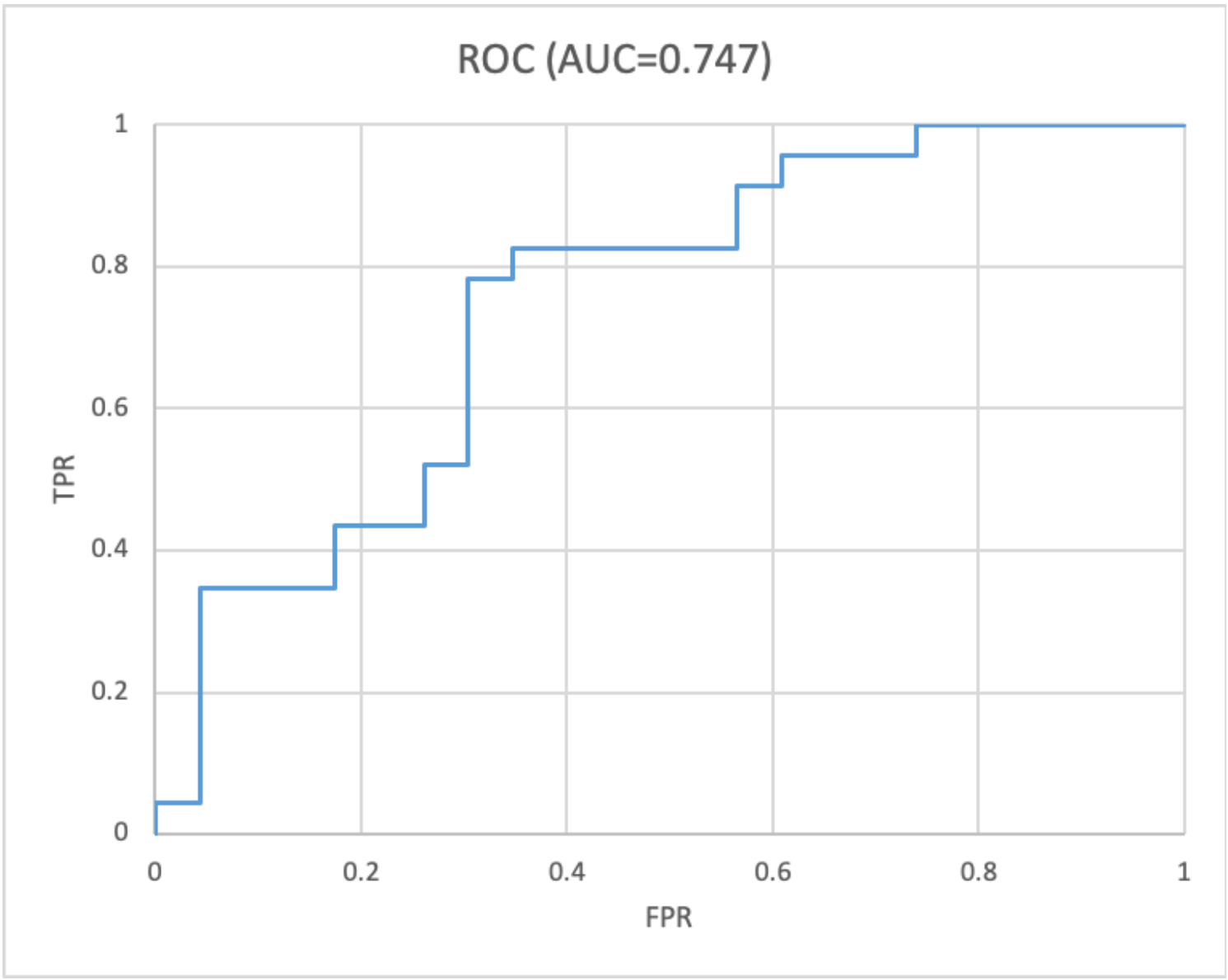

Figure 1

Receiver operating characteristic (ROC) curve for determination. The AUC was 0.746 (95% confidence interval [CI]: 0.603–0.889).