

# Prädiktion des Therapiebedarfs bei neovaskulärer AMD (nAMD) mittels annotationsbasierter Deep-Learning-Modelle

K. Rothaus<sup>1</sup>, B. Heimes-Bussmann<sup>1</sup>, S. Aydin<sup>3</sup>, M. Quassowski<sup>3</sup>, M. Ziegler<sup>1</sup>, C. Lange<sup>1,5</sup>, G. Spital<sup>1</sup>, A. Lommatzsch<sup>1,2,4</sup>, M. Gutfleisch<sup>1</sup>

<sup>1</sup>Augenzentrum am St. Franziskus-Hospital, Münster • <sup>2</sup>Zentrum für Augenheilkunde, Universität Duisburg-Essen, Essen • <sup>3</sup>Westphalia DataLab GmbH, Münster • <sup>4</sup>Achim Wessing Institut für Ophthalmologische Diagnostik, Universität Duisburg-Essen, Essen • <sup>5</sup>Klinik für Augenheilkunde, Universitätsklinikum Freiburg, Freiburg

## Einführung

Bei der nAMD-Therapie mit intravitrealen anti-VEGF Injektionen (IVI) ist die Prädiktion erneuter Krankheitsaktivität maßgeblich für die Einordnung der Patienten in einen konkreten Behandlungsplan. Ziel ist es sowohl Über- als auch Unterbehandlung zu vermeiden und die Kontrollintervalle bedarfsgerecht zu organisieren.

Anhand von Experten-Annotationen bekannter SD-OCT nAMD-Biomarkern (→ [1]) wurde ein XAI-Modell (eXplainable Artificial Intelligence) entwickelt. Hierzu trainierten wir zunächst Segmentierungsmodelle (→ [2]). Anschließend wurden weitere KI-Modelle entwickelt um drei prädiktive Zielfragen zu bearbeiten:

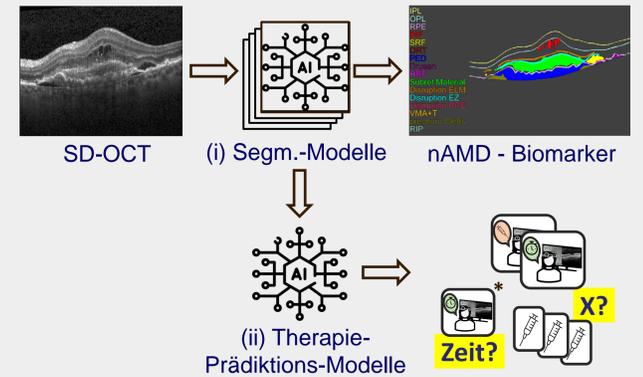
1. Aktivität der nAMD bei der nächsten Kontrolle,
2. Zeitpunkt der nächsten Aktivität und
3. Injektionsbedarf im nächsten Behandlungsjahr.

## Patienten und Methode

Es wurden 1.163 longitudinale Datensätze von 936 nAMD-Patienten (580 weiblich) aus der Sprechstunde des Augenzentrums der Jahre 2012 – 2022 analysiert, die überwiegend nach pro re nata Schema (IVAN-Protokoll) behandelt wurden. Bei einer durchschnittlichen Beobachtungsdauer von 2,8 Jahre ( $\pm 4,1$  SD), flossen 18.915 IVI- und 17.539 SD-OCT-Termine mit in die Auswertung ein. Insgesamt wurden die KI-Modelle an 49.577 SD-OCTs (HE-Spectralis®, 49 B-Scans) trainiert und „eXplainable AI“-Modelle (XAI) entwickelt.

Initial annotierten fünf Junior- und ein Seniorgrader 4.177 zufällig ausgewählte B-Scans durch Einzeichnen von drei Netzhautschichten und 17 Biomarkern (→ [1]); u.a. intra- (IRF) und subretinale Flüssigkeiten (SRF), subretinale Material, Pigmentepithelabhebung (PED) und Atrophien. Anhand dieser Annotationen wurden Segmentierungsmodelle (Abb.1 Schritt (i)) trainiert und fünf Mal fünffach kreuzvalidiert (→ [2]).

In einem zweiten Schritt wurden die Prädiktions-Modelle (Abb.1, Schritt (ii)) konzipiert und trainiert, um auf Basis der SD-OCT Daten bzw. KI-Segmentierungen den IVI-Therapiebedarf bei nAMD vorherzusagen (Experimente 1 bis 3), erneut 5 x 5-fach kreuzvalidiert.



**Abb.1:** XAI-Design: (i) Biomarker-Segmentierung (ii) Prädiktion

## Ergebnisse

### Experiment 1: IVI-Indikation bei nächster Kontrolle

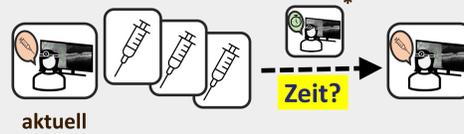


Bei der Vorhersage der nächsten Verlaufskontrolle konnte die Güte von 68,5% AUC (Area Under Curve) ohne Segmentierungen auf 73,8% gesteigert werden. Bei Wahl des „Decision-Points“ mit maximalem F1-Score wurden 74,9% der inaktiven ( $C_0$ ) und 63,2% der aktiven ( $C_1$ ) Verlaufskontrollen korrekt vorhergesagt.

Fallzahl:  $n_0 = 1.819$ ,  $n_1 = 1.184$ ,  $n_{out} = 0$

Güte: TPR = 63,2%, TNR = 74,9%, Cohens  $\kappa = 0,38$

### Experiment 2: Zeit bis zur nächsten IVI-Indikation

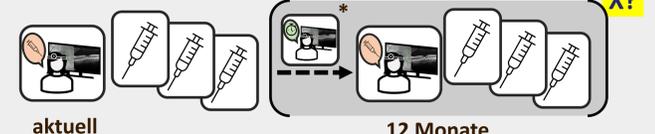


Bei der Prädiktion des Zeitintervalls  $T_{stabil}$  ohne Aktivität nach der aktuellen Serie konnte die Güte von 70,4% AUC auf 77,4% (mit Segmentierungsmodellen) gesteigert werden. Dabei wurden zwischen Fällen mit langem ( $C_0$  mit  $T_{stabil} > 12$  Wochen) und kurzem Inaktivitätsintervall ( $C_1$  mit  $T_{stabil} \leq 6$  Wochen) unterschieden.

Fallzahl:  $n_0 = 557$ ,  $n_1 = 800$ ,  $n_{out} = 857$  (mit  $6 < T_{stabil} \leq 12$ )

Güte: TPR = 69,2%, TNR = 73,6%, Cohens  $\kappa = 0,39$

### Experiment 3: Anzahl der IVI-Serien im nächsten Jahr

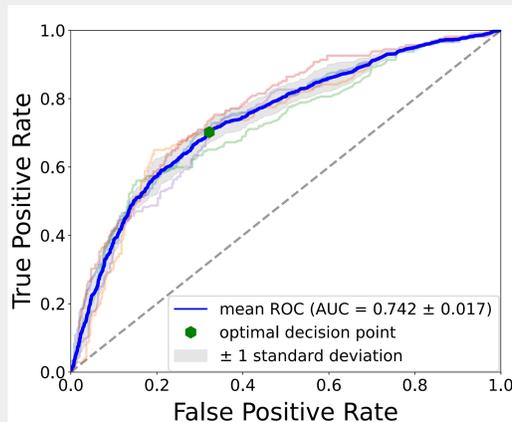


Bei Prädiktion der Injektionsfrequenz konnte bei Unterscheidung zwischen „höchstens 1 Serie“ ( $C_0$ ) und „mindestens 3 Serien“ ( $C_1$ ) die Güte von 73,2% AUC auf 78,1% AUC gesteigert werden. Bei der Ausweitung von  $C_0$  auf „ $\leq 2$  Serien“, konnten wir eine Güte von 73,9% AUC erreichen (63,2% ohne Segmentierungen).

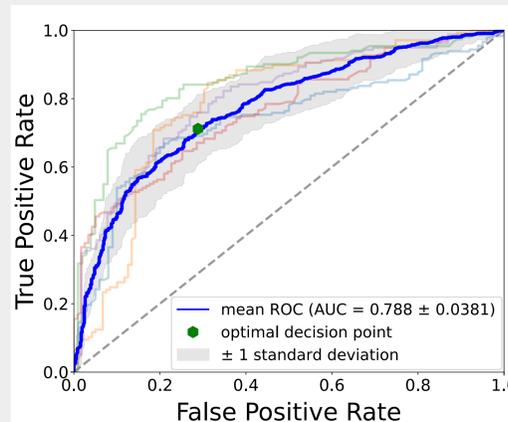
Fallzahl:  $n_0 = 312$ ,  $n_1 = 929$ ,  $n_{out} = 956$  („ $\leq 2$  Serien“)

Güte: TPR = 70,0%, TNR = 72,7%, Cohens  $\kappa = 0,34$

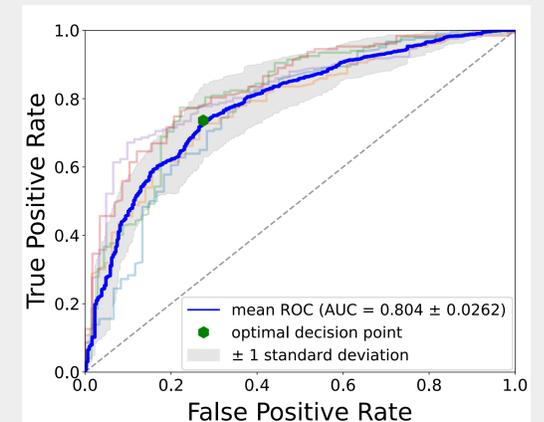
## Abbildungen



(a) ROC zur Prädiktion der Aktivität zur nächsten Kontrolle



(b) ROC zur Prädiktion einer kurzen ( $\leq 6$  Wochen) Inaktivitätsintervall



(c) ROC zur Prädiktion von einer hohen ( $\geq 3$  Serien) IVI-Frequenz

**Abb. 2:** ROCs (Receiver Operation Characteristic) zu den Experimenten 1 bis 3 für je einen Durchlauf. Die TPR (True Positive Rate) bezieht sich auf Klasse  $C_1$ . Optimaler „Decision Point“ bei max. F1-Score.

## Diskussion und Schlussfolgerungen

Es wurden innovative XAI-Modelle zur Prädiktion des Therapiebedarf bei nAMD entwickelt. Die Annotation relevanter Strukturen in SD-OCTs half die Güte der KI-Modelle (→ [3]) deutlich zu verbessern. Auch der Therapiebedarf bei der nächsten Kontrolle sowie die Prognose der nächsten Monate konnten hiermit sehr gut vorhergesagt werden. Indem durch bessere Prädiktion unnötige Arztbesuche vermieden werden, erwarten wir die Krankheitslast („burden of disease“) zu verringern. Wir sehen die Ergebnisse als wichtigen Schritt zur Individualisierung der anti-VEGF Therapie.

Durch die automatische Segmentierung erhalten die Ärzte ein Werkzeug zur Einordnung der KI-Entscheidung. Es ist denkbar, zukünftig durch Interaktivität die Einschätzung der Ärzte zu berücksichtigen sowie

„Active-Learning“ Methoden umzusetzen.

Falls sich im Rahmen einer Verlaufskontrolle eine Aktivität der nAMD in der SD-OCT zeigte, erwiesen sich die auf den Biomarker basierenden Modellen als Verbesserung gegenüber reinen SD-OCT-Modellen. Bei einer nAMD ohne Krankheitsaktivität und ohne IVI-Indikation konnten wir keine Verbesserung durch Annotationen feststellen. Dies zeigt, dass die nAMD-Biomarker zwar die Aktivität, aber nicht die Inaktivität geeignet quantifizieren können.

Unsere Ergebnisse sollen zukünftig auf Daten mit anderen Behandlungskonzepten (z.B. Treat & Extend) sowie länger wirkenden Präparaten validiert werden. Erste Experimente zu anderen SD-OCT-Herstellern und SD-OCT-Formaten waren vielversprechend.

## Literatur

- [1] Metrangola C, Donati S, Mazzola M et al.: **OCT Biomarkers in Neovascular Age-Related Macular Degeneration: A Narrative Review.** J Ophthalmol 2021; doi: 10.1155/2021/9994098
- [2] Gutfleisch M., Heimes-Bussmann B., Aydin S. et al.: **„Annotation von SD-OCT-Biomarkern bei nAMD zur Entwicklung erklärbarer KI-Modelle (XAI)“** Kongressbeitrag PFr05-05 DOG 2022
- [3] Gutfleisch M. et al.: **„Clinically applicable deep learning-based decision aids for treatment of neovascular AMD“.** Graefes Arch Clin Exp Opt. 2022, 260(7); doi: 10.1007/s00417-022-05565-1.

## Danksagung

Wir bedanken uns bei der Firma Novartis Pharma GmbH für ihre Förderung und Unterstützung dieser Arbeit sowie der Westphalia DataLab GmbH für die Zusammenarbeit.

