

A 3-Dimensional Convolutional Neural Network Model to Estimate the Glaucomatous Visual Field from Optical Coherence Tomography

Author

Makoto Koyama¹

Author affiliation

1: Minamikoyasu Eye Clinic, 2-8-30, Minamikoyasu, Kimitsu-shi, Chiba-ken, 299-1162, Japan

Corresponding Author

Makoto Koyama

Address: Minamikoyasu Eye Clinic, 2-8-30, Minamikoyasu, Kimitsu-shi, Chiba-ken, 299-1162, Japan

Telephone: (81) 439271022

E-mail: minamikoyasuganka@gmail.com

Keyword

optical coherence tomography, glaucoma, visual field, machine learning

Abstract

Purpose: To improve the accuracy of glaucomatous visual field (VF) estimations from optical coherence tomography (OCT) images, we trained a segmentation-free 3-dimensional convolutional neural network (3DCNN) model instead of the conventional 2-dimensional model.

Design: Retrospective, single-clinic all-cases survey.

Method: We investigated 3,416 patients (6,356 eyes) who underwent OCT and Humphrey Field Analyzer (HFA) 24-2 or 10-2 tests, including all types of glaucoma or suspected glaucoma. All patients containing intraocular disease with or without glaucoma and all images with a weak signal strength index, except those with VF defects due to extraocular diseases, were included. The teacher values of the VF thresholds to be learned and of the VF threshold slopes to be learned were calculated from the pointwise linear regression of the VF thresholds. The mean VF values of all measurement points were 26.5 ± 6.91 dB [mean \pm standard deviation] for HFA24-2 and 27.8 ± 8.30 for HFA10-2.

Results: The root mean square error per paired data for glaucoma cases were 3.27 ± 2.20 dB for HFA24-2 and 3.02 ± 2.28 dB for HFA10-2. The Pearson's correlation between the VF and the estimated VF in glaucoma cases were $r=0.848$ for HFA24-2 and $r=0.882$ for HFA10-2. The Pearson's correlation between the mean deviation (MD) and the estimated MD in glaucoma cases were $r=0.903$ for HFA24-2 and $r=0.923$ for HFA10-2. The Pearson's correlation between the VF slopes and the estimated VF slopes in glaucoma cases were $r=0.431$ for HFA24-2 and $r=0.438$ for HFA10-2. The Pearson's correlation between the MD slopes and the estimated MD slopes in glaucoma cases were $r=0.429$ for HFA24-2 and $r=0.383$ for HFA10-2.

Conclusion: The 3DCNN model achieved high VF estimation accuracy using OCT images.

光干渉断層計画像から緑内障の視野を推測する 3次元畳み込みニューラルネットワークモデルの構築

著者

古山 誠¹

著者の所属機関

1: 南子安眼科

住所 〒299-1162 千葉県君津市南子安2-8-30、日本

責任著者

氏名 古山 誠

住所 〒299-1162 千葉県君津市南子安2-8-30、日本

電話番号 0439-27-1022

電子メールアドレス minamikoyasuganka@gmail.com

キーワード：光干渉断層計、緑内障、視野、機械学習

要約

目的: 光干渉断層計 (OCT) 画像からの視野推定精度を向上させるため、従来の2次元モデルに代えて、セグメンテーション不要の3次元畳み込みニューラルネットワークモデル (3DCNN) により学習を行った。

対象: 単一施設の後向き全例調査。

方法: OCT に加えハンフリー視野計(HFA) 24-2もしくは10-2検査を受けた、すべてのタイプの緑内障または緑内障の疑いの患者を含む3,416人 (6,356眼) を調査した。眼外疾患による視野欠損例を除き、緑内障の有無に関わらず眼内疾患の含有例も信号強度指数が低い画像もすべて学習対象とした。学習すべき視野及び視野の傾きの教師値は、視野閾値の測定点毎の回帰直線から算出した。全測定点の視野の平均教師値は、HFA24-2 で 26.5 ± 6.91 dB [平均 \pm 標準偏差]、HFA10-2 で 27.8 ± 8.30 dB であった。

結果: 緑内障患者のペアデータあたりの二乗平均平方根誤差は、HFA24-2で 3.27 ± 2.20 dB、HFA10-2で 3.02 ± 2.28 dB であった。緑内障患者の測定点毎の視野と推測視野のピアソンの積率相関係数はHFA24-2で $r=0.848$ 、HFA10-2で $r=0.882$ であった。緑内障患者のMD値と推測MD値のピアソンの積率相関係数は HFA24-2で $r=0.903$ 、HFA10-2で $r=0.923$ であった。緑内障患者の測定点毎の視野の傾きに関し、教師値と推測値のピアソンの積率相関係数は HFA24-2で $r=0.431$ 、HFA10-2で $r=0.438$ であった。緑内障患者のMD値の傾きに関し、教師値と推測値のピアソンの積率相関係数は HFA24-2で $r=0.429$ 、HFA10-2で $r=0.383$ であった。

結論: 3DCNN モデルにより、OCT 画像からの高い視野推測精度を達成した。

緒言

緑内障による視野欠損は非可逆的な変化¹で、その進行により患者の生活の質は大きく低下する恐れ²がある。緑内障の視野欠損の状況を過小評価すると治療の遅れに繋がり、逆に過大評価すると本来不要であるはずの治療を行う原因となりうる。視野欠損の評価を正確に行うために視野検査を頻繁に行うと、視野検査は時間のかかる検査であるため、患者側及び医療側の両者の負担が大きくなってしまふ。その点、光干渉断層計(OCT)は視野検査よりも短い時間で測定でき、緑内障が進行するほど OCT で測定される網膜の内層側の厚みが薄くなる事が知られている³⁻⁶ため、OCT の画像から視野を推測する事はとても大切な事である。

OCT から視野を推測するモデルの報告は過去に多数³⁻¹²あり、そのほとんどが2次元の機械学習モデルで、症例数も限られたものであった。2次元のモデルの場合には、網膜の層毎に OCT のセグメンテーションを行う必要があるため、しばしばセグメンテーションエラーを起こす事があった。セグメンテーションエラーの有無を人間が全て確認するのはとても大変な作業になるため、合併疾患を除外して、ノイズが少なく綺麗に撮影できた OCT のみで学習する必要があり、その結果どうしても学習対象となる症例数が限られてしまう傾向にあった。

機械学習の学習精度を上げるには、大量の症例数を準備する必要がある。もし3次元の OCT 画像をそのまま3次元の機械学習モデルで学習すれば、セグメンテーションが不要なため、緑内障以外の疾患も含んで学習する事も可能で、ノイズの多い OCT も全て学習対象とする事ができ、大量の症例数が確保でき、同時に黄斑上膜等の合併疾患のある症例やノイズ混じりの OCT の推測精度の向上も期待できる。そのように考え、3次元の機械学習モデルの中でも推測精度の優れた、3次元畳み込みニューラルネットワーク(3DCNN)を使用して、教師あり学習による機械学習モデルを作成した。

学習対象をどの程度まで広げるのが最も良い結果が出るのかを検証するため、緑内障以外の眼内疾患も全て含んで学習した群と、眼内疾患の合併症を除外して緑内障のみの症例で学習した群の2つに分け、結果を比較した。さらに、OCT 画像の信号強度指数 (SSI) が低い (=ノイズの多い) OCT を含んで学習した群と、SSI が低い OCT を除外して学習した群の2つも分け、結果を比較した。

また、OCT 撮影時の視野の推測だけではなく、視野の進行速度も同時に推測できれば有用と考え、それぞれ1枚ずつの OCT からの視野の進行速度も同時に学習した。

研究方法

後ろ向きの単一クリニックの全例調査である。南子安眼科（当院）の院内に倫理委員会を設立し承認を行い、研究情報を院内及びホームページに開示して拒否の権利を保証し、ヘルシンキ宣言に則り、人を対象とする医学系研究の倫理指針に準拠し、個人情報保護に関する法律について

てのガイドライン Q&A2-15（利用目的による制限の例外）に準拠し実施した。

症例選択

当院を開院してから現在まで（2011年3月8日～2022年8月4日）の単一施設の後ろ向き全例調査である。

OCTの黄斑マップに加え、ハンフリー視野計(HFA) 24-2/30-2もしくは10-2を1回以上施行した症例を学習対象にした。疾患の診断は全て1人の医師（私）が行い、同一の医師がカルテを元に後ろ向きに症例を数え、重複疾患も全て加算して記載した。眼底検査は全例 MaxField78D レンズを細隙灯顕微鏡で使用し、無散瞳での検査の症例も含まれる。偽陽性33%、偽陰性33%、固視不良33%をカットオフ値とし、カットオフ値以上の HFA は除外した。HFA30-2は周辺を削って HFA24-2のデータとして学習対象にした。

緑内障以外の症例も全て含んで学習した方が推測性能の向上が認められるかどうかを確認するため、表1の眼外疾患を除外した全ての症例を対象に、学習時の条件により以下の2群に分けた。眼内疾患含有群 は、学習時に表1の眼内疾患を含め、緑内障の有無に関わらず緑内障以外の疾患も含めて全ての HFA と全ての OCT の3次元黄斑マップ画像を元に機械学習を行った。眼内疾患除外群 は、学習時に表1の眼内疾患を除外し、緑内障・緑内障疑い・視神経乳頭陥凹拡大以外の疾患を合併している症例を全て除外して学習した。

いずれの群も学習時には緑内障・緑内障疑い・視神経乳頭陥凹拡大は区別せず、表1に記載の無い、白内障、後発白内障、強度近視、角膜混濁、点状表層角膜炎、翼状片等の前眼部・中間透光体病変の合併例は除外せず全ての群で学習対象とした。いずれの群も、緑内障の病型の区別は行っておらず、隅角の所見による区分、ぶどう膜炎合併の有無による区分、急性・慢性の区分、眼圧による区分はいずれも行っていない。強度近視合併の緑内障及び後部ぶどう腫合併の緑内障の場合も除外せず、全ての群で学習対象及び評価対象とした。近視性視神経症による視野欠損は緑内障による視野欠損との鑑別が困難な場合が多いと考え、今回は両者の区別をせずに全て緑内障として学習対象及び評価対象とした。但し、強度近視による網脈絡膜萎縮もしくは網膜剥離等の合併疾患を認めた場合は、表1の眼内疾患に区分した。

眼外の合併疾患も学習対象とすると、本来の目的である緑内障の推測精度に悪影響が出てしまう恐れがあるため、表1の眼外疾患の症例は全ての群の学習対象から除外した。

解析の母集団が異なるとモデルの精度比較が困難になるため、学習対象と評価方法は切り分けて考え、それぞれの学習対象で学習したモデルに対し、最後に同一条件のテストデータで評価した。即ち、眼内疾患含有群と眼内疾患除外群は学習対象が異なるが、精度比較のため評価対象を同一に揃え、例えば両群とも「緑内障以外の合併疾患症例を除外した同一群に対して評価」のように行った。評価対象を変えた時にも単回の学習で対応できるように、モデルからの出力時に全

ての症例からの推測を保存し、後で評価対象の条件に応じて計算するという方法を取った。

ノイズ混じりの OCT から学習した方が良い推測性能が得られるかどうか確認するため、全ての SSI を含んで学習した群（全 SSI 群）と、SSI が 7 未満の症例を除外して学習した群（SSI 制限群）に分けてモデルを作成し、最後に評価対象を同一に揃え比較を行った。SSI が極めて低い症例の緑内障の診断は、白内障や後発白内障が進行して SSI が低くなった場合がほとんどであるため、その症例の眼底が見えた過去の時点の所見、もしくは手術後に眼底が見えるようになった後の所見を元に診断を行った。初診時から終診時まで一貫して白内障が強く眼底が見えない症例は、もともと視野検査を施行していないため、今回の研究対象には含まれていない。

当院では、検診で視神経乳頭陥凹拡大を指摘された方に OCT 及び HFA24-2 を行う事も多かったため、その場合の多くは正常な視野であり、モデルからの推測が比較的容易であると推測される。そのような症例も全て含んで評価を行うと、本来の推測性能よりも良い結果が出てしまう恐れがあるため、学習時にはいずれの群も Anderson-Patella の基準(パターン偏差で $p < 5\%$ の点が 3 つ以上隣接しそのうち 1 点以上が $p < 1\%$ 、パターン標準偏差 $< 5\%$ 、緑内障半視野テストが正常範囲外のいずれか)¹³に合致しない症例も全て対象に含めたが、評価時には眼毎に最終 HFA24-2 が Anderson-Patella の基準に合致する HFA24-2 のみを評価した。緑内障の発症初期の場合、Anderson-Patella の基準に後から合致する症例もあるため、そのような症例は途中経過も含め評価対象に含めた方が良くと考え、途中の視野が例え Anderson-Patella の基準に該当していない場合でも、最終 HFA24-2 が Anderson-Patella の基準に該当している場合には眼毎に過去に遡って全て評価対象とした。

尚、当院での HFA10-2 は、OCT 所見から見て中心視野が欠損している可能性が考えられる症例に限り行っており、また HFA24-2 が Anderson-Patella の基準に該当していない場合でも HFA10-2 の中心視野欠損を来す症例も存在するため、HFA10-2 は HFA24-2 の Anderson-Patella の基準とは無関係に評価した。

学習対象に含めるかどうか、 2群に分けて検討した疾患 (眼内疾患)	全ての群で学習対象から除外した疾患 (眼外疾患)
<ul style="list-style-type: none"> ・黄斑上膜 103眼 ・糖尿病黄斑症(severe NPDR 以上) 81眼 ・加齢黄斑変性 or 近視性黄斑症 57眼 ・視神経(部分)低形成 44眼 ・網膜静脈(分枝)閉塞症 46眼 ・網膜動脈(分枝)閉塞症 23眼 ・網脈絡膜萎縮 33眼 ・網膜色素変性症 13眼 ・黄斑円孔 6眼 ・網膜剥離術後 6眼 ・その他眼内疾患の合併例 33眼 	<ul style="list-style-type: none"> ・両耳側半盲 ・同名1/4盲 ・パーキンソン病 ・認知症 ・アルツハイマー病 ・レンズの枠の暗点のある症例 ・上眼瞼の視野欠損のある症例 ・心因性視野障害 ・その他、眼底・OCT 所見と矛盾する視野の症例

表1 学習対象に含めるかどうかの疾患一覧。眼内疾患を合計すると除外された症例数と計算が合わない理由は、合併疾患の有無を数えた後に、HFAの33%のカットオフ値で全てのペアデータが除外されてしまった症例及び、重複疾患も含まれるためである。

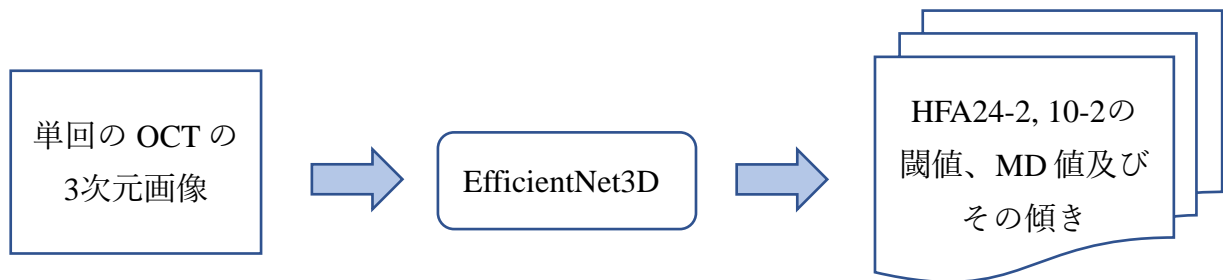


図1 今回のモデルのイメージ図。単回の OCT の黄斑マップ 3次元画像の情報を3DCNN モデルの EfficientNet3D へ与え、HFA24-2、10-2の閾値及びその MD 値とその傾きを同時に出力した。

視野検査の結果の取得方法

当院では、HFA からの出力は全て株式会社ビーラインの BeeFiles for HFA へ保存しており、HFA データの抽出を全例 HFA から直接手作業で行うと極めて長い時間がかかるため、BeeFiles for HFA から間接的に行った。自作のマクロを作成して1症例ずつ BeeFiles for HFA から視野の export を繰り返し、text データとして出力された全てのファイルを元に、Python を利用したプログラムを使用してデータを抽出し、HFA30-2・24-2ではパターン偏差・パターン標準偏差・緑内障半視野テストから Anderson-Patella の基準に該当するかどうか自動的に計算し、データベースへ登録した。

OCT の画像の取得方法

OCT は Nidek 社の RS-3000もしくは RS-3000 Advance2を使用した。2011年3月8日～2022年8月4日の対象期間の中で、2022年4月21日～2022年5月12日及び2022年6月16日～2022年8月4日の期間は RS-3000 Advance2を使用し、それ以外の期間は RS-3000を使用した。撮影条件は、RS-3000及び RS-3000 Advance2の設定を統一し、黄斑マップの wide スキャン9mm の X-Y 方向、幅9mm、高さ9mm、ScanPointA=512、ScanPointB=128、ScanCenterX=399、ScanCenterY=299、ScanWidth1=30.0、and ScanWidth2=30.0に設定し、他の撮影条件の画像は破棄した。3次元黄斑マップ+3次元乳頭マップの両方を同時に学習対象とした方が良い結果が出る可能性も考えられたが、3次元乳頭マップは当院での症例数が少ないため採用を断念し、3次元黄斑マップのみを採用した。

普段の診療では2次元に圧縮された OCT 画像を利用する機会が多いが、実際の画像情報は3次元のまま内部に保存されている。推測性能向上のため、今回の3DCNN モデルではその3次元 OCT 画像をそのまま学習に使用する必要があった。但しその保存場所は不明であり、OCT 内部のデータベースも非公開のため、患者 ID 及び撮影日時から3次元 OCT 画像の保存場所を特定するために以下の作業を行った。Tesseract-OCR を使用した Python プログラムを作成し、OCT から jpeg 形式でエクスポートされた画像を元に、撮影日時の部分を自動で切り取って OCR 解析してテキストデータへ変更し、同時に SSI や focus(近視の度に類似)の値も同様に読み取り、データベースに登録した。過去の jpeg エクスポート画像は OCT のソフトウェアのバージョンによってテキストの部位が異なるので、それぞれのソフトウェアのバージョンに対応した部位で自動的に切り取りが行えるようにプログラムを作成した。3次元の非圧縮 OCT 画像のディレクトリに存在する xml ファイルに記録されている撮影日時と、jpeg エクスポート画像の撮影日時が一致する物を自動的に探す事により、患者 ID 及び撮影日時データと3次元の非圧縮画像を過去に遡り全て紐付けしてデータベースに登録した。

OCT と視野検査のペアデータの作成方法

「単回の OCT の 3 次元画像」を入力データとし、「HFA24-2・10-2のそれぞれの閾値と Mean Deviation(MD 値)及びその傾き」を教師データとした。そのペアデータの作成方法は以下の通りである。

視野検査に関し、当院では患者側の負担¹⁴を考え、緑内障患者の視野検査は HFA24-2, 10-2の SITA-Fast で行っていた。SITA-FAST と SITA-Standard は、既報では視野の進行判定能力にはほとんど差がないとの事^{15,16}であるが、SITA-FAST は測定の際ばらつきが大きいため、そのままでは機械学習の教師値としては不適切と考え、眼毎に測定点毎の実測閾値の時間軸に対する通常の最小二乗法による回帰直線を作成し、OCT 撮影時の回帰直線からの値を視野閾値の教師値とした。但し、測定回数が少ない場合の回帰直線を延長するとばらつきが大きくなってしまいうため、測定回数×6ヶ月間を有効期限とし、OCT 撮影時と直近の HFA 測定時の期間（絶対値）が有効期限を超えた場合には OCT データを破棄した。また、緑内障は非可逆的な疾患のため、回帰直線の傾きが 0 以上の場合には測定誤差と考えて傾きを 0 とした。視野閾値の教師値は 0dB を最小値、35dB を最大値とした。MD 値も同様に回帰直線から計算し、MD 値に対しては最小値及び最大値の設定は行わなかった。

視野の進行速度に関して、MD 値の傾き（MD スロープ）は「スピアマンの順位相関係数から眼毎に計算した p 値 < 5%」の場合に最小二乗法による時間軸に対する回帰直線の傾きを採用した。視野の測定点毎の傾きは、単純に「測定点毎の p 値 < 5%」のみで採用すると偶然 5%未満として採用された本来の傾きとは異なる大量の外れ値が出てしまう恐れがあるため、「MD 値の傾きの p 値 < 5%」且つ「測定点毎の p 値 < 5%」の場合に採用した。

尚、線維柱帯切除術・切開術を受けた症例や緑内障点眼治療を行っている症例も全て学習対象としたため、未治療での進行速度というのは不明で、未治療での自然経過よりも進行速度が遅く判定されるはずである。

MD 値の傾き及び視野の傾きの学習を SSI が低い OCT 画像も含めて行くと学習毎のばらつきが大きくなる恐れがあるため、MD 値の傾き及び視野の傾きの学習は SSI 7 以上の OCT 画像に限定した。

閾値及び傾きには多くの欠損値があるため、その部位はマスク値を 1 とし、学習時の loss の計算の時に(1 - マスク値)を掛ける事により学習に影響を与えないようにし、検証時にも計算から除外した。

視野の進行速度の補正方法

以下の2種類の進行速度の補正操作を、予め学習前に全てのペアデータに関して行った。

1種類目の進行速度の補正操作は正の傾きの補正に関連するものである。

眼毎の真の視野の進行速度を測定する事は不可能であるが、視野検査を繰り返して回帰直線を作成する事で真の視野の進行速度に近い値を得る事ができる可能性が高くなるはずである。その計算上の回帰直線の傾きは、測定誤差により真の視野の進行速度よりも大きくなる事もあれば小さくなる事もあると考えられる。真の視野の進行速度がもともと遅い場合、回帰直線の傾きが正の方向にずれた場合には計算上の傾きが0より大きくなる事がしばしばある。本来、緑内障による視野欠損は非可逆的な変化のため、傾きが正の値になる事は無いはずである。傾きが正の場合に単純に傾きを0に置き換えた場合、正の傾きの測定誤差のみを補正した事になり、負の傾きの測定誤差は判別不能なため、負の傾きの測定誤差によって実際の臨床よりも進行速度の傾きが負の方向に大きくなる恐れがある。その解決策として、HFA24-2及びHFA10-2のそれぞれに関し、「MD値の傾き」及び「視野の測定点毎の傾き」の両者とも、時間軸に対する最小二乗法による回帰直線の傾きが正の場合に傾きを0と置き換えた上で、「傾きが正の場合も受け入れた時の傾きの平均」と、「傾きが正の場合に傾きを0と置き換えた時の傾きの平均」の比を全てのペアデータに対して掛ける事により、誤差の補正を行った。正の傾きの補正に関連する進行速度の最終的な補正值の比率は、HFA24-2の視野の傾きは0.883倍、HFA24-2のMD値の傾きは0.822倍、HFA10-2の視野の傾きは0.933倍、HFA10-2のMD値の傾きは0.881倍となった。

2種類目の進行速度の補正操作はSSI低下に関連するものである。

例えば中には患者が手術を希望せずに白内障や後発白内障が進行する事により見かけ上の視野欠損が進行してしまう症例もあり、そのような症例を含んでそのまま学習すると、本来の緑内障による進行速度よりも早い進行速度で学習してしまう恐れがある。その問題を解決するため、以下の通りにHFA24-2及びHFA10-2の進行速度の補正を行った。一般的には、白内障や後発白内障が進行すると眼内へ光が透過し難くなるためにOCTのSSIの数値が低下^{17,18}し、またOCTのSSIの数値が低下するとMD値や視野閾値の低下が認められる(図12~図15)。年齢と共にSSIは低下する傾向にあるため(図16~図17)、SSIが有意に減少した場合との差分を計算し、その差分を元にして進行速度の補正を行った。その具体的な計算方法は以下の通りである。今回の研究での全てのペアデータに関し、「SSIに対するMD値(dB/SSI)」の最小二乗法による回帰直線の傾き及び、「SSIに対する平均視野(dB/SSI)」の最小二乗法による回帰直線の傾きを算出する事により、SSIが1低下する度に平均どの程度MD値及び平均視野が悪化するのかを計算した。次に、加齢によりSSIの数値は低下するため、全てのペアデータに関し眼毎に「年齢に対するSSI (SSI/year)」

の最小二乗法による回帰直線の傾きを算出する事により、年齢が1歳増える度に平均で SSI がどの程度減少するのかを計算した。その上で、眼毎に「時間軸に対する SSI (SSI/year)」の最小二乗法による回帰直線の傾きの「時間軸に対する SSI のスピアマンの順位相関係数から計算した p 値<5%」を有意と判定し、有意な場合にはその傾き(SSI/year)と上記の「全てのペアデータの年齢に対する SSI (SSI/year)」を比較し、「全てのペアデータの年齢に対する SSI (SSI/year)」よりもその傾き(SSI/year)の絶対値が大きい場合には、「眼毎の時間軸に対する SSI の傾き(SSI/year)」から「全てのペアデータの年齢に対する SSI (SSI/year)」を引く事により平均的な SSI の低下速度との差分を算出し、その値に「SSI に対する MD 値(dB/SSI)」もしくは「SSI に対する平均視野(dB/SSI)」の値を掛ける事により、SSI の減少による分の「MD 値の傾き(dB/year)」もしくは SSI の減少による分の「視野の傾き(dB/year)」が得られ、次にその値を元の「MD 値の傾き(dB/year)」もしくは元の「測定点毎の視野の傾き(dB/year)」に対し減算する事により、SSIの減少による見かけ上の傾きへの影響を打ち消した。その結果、例えば白内障や後発白内障が進行した事による見かけ上の視野欠損進行をある程度打ち消した進行速度が算出されるため、補正前と比べてその分遅い進行速度となり、最も知りたい緑内障の進行による視野欠損の速度により近いものになるはずである。尚、傾きが0よりも大きくなる場合には0とし、「時間軸に対する SSI のスピアマンの順位相関係数から計算した p 値<5%」が得られない場合には、SSIが低下しているかどうか不明であるため上述の補正を行っていない。そのため、SSI減少による進行速度への影響を完全に打ち消す事はできず、補正を行った眼と補正を行っていない眼の両方がペアデータに存在する事になる。SSI低下に関連する進行速度の補正操作による全体への影響を計算すると、最終的な補正值の平均値の比率は、HFA24-2の視野の傾きは0.974倍、HFA24-2のMD値の傾きは0.973倍、HFA10-2の視野の傾きは0.979倍、HFA10-2のMD値の傾きは0.970倍となった。

尚、視野の測定点毎の傾きはばらつきが大きいため、上記の2種類の進行速度の補正とも、最終的な補正の比率は全ての測定点の補正の比率の平均値から算出した。表2, 表3に記載の値は上記の2種類の進行速度の補正後の値である。

Anderson-Patella	Anderson-Patella の基準に該当しない症例も全て学習対象			
学習対象疾患	眼内疾患含有	眼内疾患除外	眼内疾患含有	眼内疾患除外
SSI の条件	全 SSI		SSI を7以上に制限	
ペアデータ総数	41,420	38,368	36,066	33,660
全患者数	3,416	3,274	3,286	3,156
内 HFA24-2の患者数	3,324	3,185	3,201	3,073
内 HFA10-2の患者数	1,241	1,171	1,208	1,142
全眼数	6,356	5,961	6,011	5,658
内 HFA24-2の眼数	6,153	5,778	5,818	5,484
内 HFA10-2の眼数	2,267	2,079	2,177	2,001
HFA24-2の平均閾値(dB)	26.5±6.91	26.7±6.68	26.8±6.74	27.0±6.53
HFA10-2の平均閾値(dB)	27.8±8.30	28.1±8.13	28.1±8.16	28.4±8.02
HFA24-2の閾値スロープ 平均(dB/年)	-0.816±1.03	-0.783±1.00	-0.816±1.03	-0.783±1.00
HFA10-2の閾値スロープ 平均(dB/年)	-1.22±1.34	-1.18±1.33	-1.22±1.34	-1.18±1.33
HFA24-2の 平均 MD 値(dB)	-3.59±5.23	-3.34±5.04	-3.34±5.05	-3.12±4.89
HFA10-2の 平均 MD 値(dB)	-4.54±5.69	-4.31±5.54	-4.27±5.54	-4.08±5.42
HFA24-2の 平均 MD スロープ (dB/年)	-0.312±0.418	-0.294±0.411	-0.312±0.418	-0.294±0.411
HFA10-2の 平均 MD スロープ(dB/年)	-0.482±0.508	-0.454±0.460	-0.482±0.508	-0.454±0.46
平均 SSI	7.93±1.34	7.97±1.33	8.29±0.97	8.31±0.972
focus(≒屈折の度)(D)	-2.55±3.79	-2.62±3.82	-2.56±3.75	-2.63±3.77
年齢(歳)	63.7±13.1	63.2±13.1	62.8±13.1	62.3±13.1
HFA24-2の閾値スロープの 教師値の採用率 (%)	5.52	4.91	5.52	4.91
HFA10-2の閾値スロープの 教師値の採用率 (%)	4.00	3.56	4.00	3.56
HFA24-2の MD スロープの 教師値の採用率 (%)	20.4	18.4	20.4	18.4
HFA24-2の MD スロープの 教師値の採用率 (%)	13.5	12.3	13.5	12.3

表2 HFA24-2, HFA10-2の学習対象の母集団及び HFA10-2の評価対象の母集団。±は標準偏差。教師値の採用率の計算は、HFA24-2 全体もしくは 10-2 全体のそれぞれを分母として行った。

Anderson-Patella	最終 HFA24-2が Anderson-Patella の基準に 該当する症例のみ評価対象			
評価対象疾患	眼内疾患含有	眼内疾患除外	眼内疾患含有	眼内疾患除外
SSI 条件	全 SSI		SSI を7以上に制限	
HFA24-2ペアデータ数	27,712	25,090	23,678	21,633
HFA24-2の患者数	2,407	2,234	2,290	2,130
HFA24-2の眼数	3,569	3,220	3,340	3,030
HFA24-2の平均閾値 (dB)	25.1±7.70	25.3±7.52	25.4±7.59	25.6±7.42
HFA24-2の閾値スロープ 平均(dB/年)	-0.922±1.06	-0.891±1.04	-0.922±1.06	-0.891±1.04
HFA24-2の 平均 MD 値(dB)	-4.93±5.72	-4.67±5.58	-4.67±5.58	-4.45±5.46
HFA24-2の 平均 MD スロープ(dB/年)	-0.393±0.441	-0.375±0.437	-0.393±0.441	-0.375±0.437
平均 SSI	7.83±1.35	7.87±1.34	8.23±0.962	8.25±0.965
focus(≒屈折の度)(D)	-2.53±3.9	-2.62±3.94	-2.56±3.86	-2.63±3.89
年齢(歳)	65.6±12.3	65.1±12.3	64.7±12.5	64.2±12.4
HFA24-2の閾値スロープの 教師値の採用率 (%)	4.78	4.20	4.78	4.20
HFA24-2の MD スロープの 教師値の採用率 (%)	15.8	14.0	15.8	14.0

表3 HFA24-2 の評価対象の母集団。教師値の採用率の計算は、HFA24-2 全体を分母として行った。

データセットの振り分け方法

全てを10分割交差検証で学習するためには極めて長い時間が必要なため、今回は5分割交差検証及び10分割交差検証の両方を用いた。5分割交差検証は複数のモデルの推測性能を比較するためのみに使用し、複数の5分割交差検証の中で最も良い結果が得られた1つのモデルの推測性能を厳密に調べるため、最良の学習条件に対し10分割交差検証を用いて再度モデルを作成し学習を行った。患者毎に全症例をランダムに等分し、5分割交差検証では学習群：検証群兼テスト群を4：1、10分割交差検証では学習群：検証群：テスト群を8：1：1に振り分けた。モデル間の比較のため、ランダムに等分する際に使用する乱数のシード値は各モデルで共通の値を採用した。

患者によってペアデータ数が多い症例もあれば、逆に少ない症例もあり、単純に全てランダムに患者毎に振り分けた場合には分割交差検証の群間でのペアデータ数が偏り、学習の妨げとなる可能性があるため、以下のように振り分けた。まず、患者毎に HFA24-2、10-2のそれぞれの時系列に対する MD 値の「スピアマンの順位相関係数から計算した p 値 $< 5\%$ 」が左右眼の少なくともどちらかで該当するか、それとも両眼とも該当しないかで患者を分けた。即ち、「少なくとも片眼の HFA24-2の MD 値の p 値 $< 5\%$ 」を満たす群と満たさない群の 2 群に分割し、さらにそれぞれ「少なくとも片眼の HFA10-2の MD 値の p 値 $< 5\%$ 」を満たす群と満たさない群に分割し、合計 4 群に分割した。それぞれの分割した 4 群に対し、ランダムに 5 分割もしくは 10 分割し、最終的にその 4 群からの患者をそれぞれ全て加える事により、均等に 5 分割交差検証用もしくは 10 分割交差検証用に患者を全て振り分けた。

表 1 の症例の除外は患者毎ではなく眼毎であるため、5 分割交差検証用もしくは 10 分割交差検証用の患者の振り分けを行った後に、眼毎に眼内疾患及び眼外疾患の除外を行った。その際、両眼とも除外される場合もあるため、最終的な 5 分割交差検証用もしくは 10 分割交差検証用の患者数は各群で必ずしも一致せず、また患者によってペアデータ数が異なるために各群でペアデータ数も一致しないが、上述の方法により、完全にランダムに振り分けた場合よりも均等に近くなる可能性が高いはずである。

学習群では OCT 画像から教師データを推測するように学習し、テスト群では OCT 画像からの推測データと正解データから精度を計算し、群をずらして 10 回繰り返して最終結果を算出した。その際、それぞれの交差検証にて検証群による早期終了を使用し、検証群での最も良いモデルを読み込んでテスト群から推測し判定した。5 分割交差検証では早期終了を使用しないで検証群無しで学習しても良いが、その場合には学習のばらつきが大きくなり、学習条件を変えた時の比較が困難になってしまうため、敢えて検証群とテスト群を兼用とした。その結果、実際よりも良い結果が出てしまう恐れがあるため、最終的なモデルの精度を算出する時には、最良の値が得られた同一学習条件にて 10 分割交差検証のモデルを再度作成し直して検証群とテスト群を分けて検証した。

MD 値及び傾きを加えて同時に学習すると、学習毎のばらつきが大きくなってしまい、どの学習条件が最適かの推測が困難になってしまうため、5 分割交差検証では HFA24-2・10-2 閾値のみを学習対象とし、その後の 10 分割交差検証では HFA24-2・10-2 の MD 値及び傾きも学習対象とした。

尚、上記の 5 分割交差検証及び 10 分割交差検証の振り分けはいずれも患者毎であるため、それぞれの群に同時に同一の患者が存在する事は無い。

視野の教師値の処理方法

視野の閾値の教師値は0dB から35dB と定義したため、35で割る事により正規化して0から1の範囲内に収める事が可能であるが、視野の進行速度の絶対値上限は母集団によって異なるため、0から1の範囲内への正規化ではなく、平均値0, 標準偏差1となる標準化が望ましいと考えた。1つのモデルで同時に学習する際、同時に標準化と正規化の教師値が混在すると、スケールの違いにより誤差逆伝播法時の学習の妨げとなる恐れがあるため、教師値は全て標準化で統一した。即ち、HFA10-2・24-2のそれぞれの学習群に対し、HFA の閾値、MD 値、HFA の閾値の傾き、MD スロープの標準偏差及び平均値を計算して標準化を行った。その際、検証群及びテスト群の教師値はモデルからは不明であるため、10分割交差検証毎もしくは5分割交差検証毎に、学習群・検証群・テスト群とも、学習群から算出した平均値及び標準偏差を使用して標準化の計算を行った。そのため、検証群及びテスト群に関しては平均値0, 標準偏差1とはならないが、それに近い値となるはずである。最終結果を算出する時にはいずれも逆の計算を行い、元の数値に戻した。

OCT データの処理方法、モデルの学習方法

今回のモデルには、PyTorch の3DCNN モデルの1つである EfficientNet3D-b0¹⁹を使用した。EfficientNet²⁰は2次元の CNN の中でも特に高い推測性能と高い学習効率が特徴のモデルで、今回使用した EfficientNet3D はその3次元版である。EfficientNet3D にも様々な大きさのモデルがあり、学習効率の面から今回はその中で最も小さい EfficientNet3D-b0を使用した。

元々512 × 512 × 128 pixel × 8bit グレースケールの3次元 OCT 画像を、224 × 224 × 128 pixel × 8bit グレースケール画像へ Lanczos 法で変換し、その値が0から1の範囲に収まるように256で割る事により正規化して EfficientNet3D-b0へ入力した。学習毎に画像を変換すると学習時間が極めて長くなるため、予め変換した Tensor イメージを local の高速なストレージに保存して読み出すようにした。過学習を抑えるため、EfficientNet3D-b0の head 部分を入れ替え、モデル内部の出力直前の1280次元からの元の dropout20%に加え、dropout をさらに30%追加した。dropout の後は隠れ層無しの linear の全層結合で出力した。モデルからの出力は、10分割交差検証では244次元の「HFA24-2・10-2のそれぞれの、閾値及び傾き及び、MD 値及び傾き」であり、5分割交差検証では120次元の「HFA24-2・10-2のそれぞれの閾値」である。バッチサイズは4、学習率は6e-4から1e-3まで3epoch かけて漸増して、その後5epoch かけて6e-4まで漸減した。Optimizer は Adam を使用し、Weight Decay は非使用である。転移学習データは見当たらなかったため使用していない。

反転方法

効率的に学習するため、左眼の OCT データは水平反転し、擬似的に右眼のデータとして学習し、それに対応した教師値も水平反転を行った。学習データを2倍に水増しするため、OCT データは

垂直反転を使用し、それに対応した教師値も垂直反転を行った。垂直反転は学習時のみではなく検証時にも用い、垂直反転した値と垂直反転していない値の平均的を取り最終出力とした。表2, 表3は水増し前の値である。

計算方法

テスト群からの推測データと正解データから、ペアデータ毎に測定点毎の二乗平均平方根誤差 (Root Mean Square Error; RMSE) を算出した。既報では先に平方根を取りその後に全ての症例の平均値を計算している報告^{4,7,10}が多いため、今回は既報に合わせ、先に平方根を取り計算した。つまり、ペアデータ毎に

$$RMSE = \sqrt{\frac{1}{m} \sum_{d=1}^m (\hat{Y}_d - Y_d)^2} \text{ とした。}$$

上記の \hat{Y} は推測データ、 Y は正解データ、 m は測定点の総数、 d は測定点の番号である。HFA24-2の閾値及び閾値の傾きの場合には $m = 52$ 、HFA10-2の閾値及び閾値の傾きの場合には $m = 68$ 、HFA24-2・HFA10-2のMD値及びMD値の傾きの場合には $m = 1$ となる。

RMSE同様、テスト群からの推測データと正解データからペアデータ毎に平均絶対誤差 (Mean Absolute Error; MAE) を算出した。閾値及び閾値の傾きのMAEは2種類計算し、測定点毎のMAEと、視野全体の平均のMAEのそれぞれを計算した。つまり、ペアデータ毎に

$$\text{測定点毎の MAE} = \frac{1}{mn} \sum_{i=1}^n \sum_{d=1}^m |\hat{Y}_{di} - Y_{di}| \text{ とし、}$$

$$\text{視野全体の平均の MAE} = \frac{1}{mn} \sum_{i=1}^n \left| \sum_{d=1}^m \hat{Y}_{di} - \sum_{d=1}^m Y_{di} \right| \text{ とした。}$$

上記の n はペアデータの総数、 i はペアデータの番号である。MD値及びMD値の傾きの場合、もともとMD値が視野全体の指標のため、上記のどちらで計算しても $m = 1$ から同一のMAEとなる。

計算は全てPythonで行った。スピアマンの順位相関係数、ピアソンの積率相関係数はPythonのSciPyを使用し、RMSE及びMAEはNumPyを使用した。SciPy及びNumPyは世界的に最も広く使用されている数値解析及び数値計算のパッケージである。

結果

学習条件及び評価条件毎の5分割交差検証の結果一覧を表4に示す。

モデル名				モデル1	モデル2	モデル3	モデル4
学習時 条件	学習時に SSI を7以上に制限したか			全 SSI		SSI7以上	
	学習時に眼内疾患を対象に含めたか			含有	除外	含有	除外
HFA24-2 評価時 条件	眼内疾患 除外	SSI 7以上	測定点毎の ピアソンの 積率相関係数	r=0.853 p<0.001	r=0.851 p<0.001	r=0.848 p<0.001	r=0.849 p<0.001
			RMSE(dB)	3.23 ±2.17	3.26 ±2.14	3.28 ±2.18	3.28 ±2.18
	眼内疾患 含有	SSI 7未満	RMSE(dB)	3.73 ±2.31	3.80 ±2.39	3.90 ±2.42	3.90 ±2.43
			RMSE(dB)	3.35 ±2.24	3.39 ±2.23	3.40 ±2.25	3.42 ±2.27
HFA10-2 評価時 条件	眼内疾患 除外	SSI 7以上	測定点毎の ピアソンの 積率相関係数	r=0.881 p<0.001	r=0.880 p<0.001	r=0.881 p<0.001	r=0.879 p<0.001
			RMSE(dB)	3.01 ±2.33	3.05 ±2.32	3.04 ±2.29	3.05 ±2.33
	眼内疾患 含有	SSI 7未満	RMSE(dB)	3.73 ±2.75	3.76 ±2.82	3.87 ±2.85	3.83 ±2.81
			RMSE(dB)	3.11 ±2.41	3.15 ±2.40	3.13 ±2.37	3.15 ±2.41

表4 5分割交差検証の学習結果。学習時及び評価時の条件を変え、眼内疾患の除外の有無及び信号強度指数(SSI)の制限の有無により比較した。二乗平均平方根誤差 (RMSE) はペアデータ毎に計算した。RMSE は小さいほど良く、ピアソンの r は大きいほど良い。太字は最良値。HFA24-2 は最終視野が Anderson-Patella の基準に該当する眼のみを評価対象とした。

モデル間の精度を比較するための5分割交差検証の結果(表4)、SSIを7以上に制限して学習したモデル(モデル3・4)よりも、全てのSSIで学習したモデル(モデル1・2)の方がRMSEは低く、特にSSI7未満の評価対象に対してのRMSEの差が大きかった。即ち、ノイズ混じりのOCTを含んで学習した方が、ノイズ混じりのOCTからの推測性能が良い傾向が見られた。

モデル1はモデル2に比べ、HFA24-2・10-2とも、RMSEが低くピアソンのrも大きかったが、その差はわずかであった。モデル3とモデル4の差はほとんど認められなかった。

以上を総合して評価すると、モデル1が概ね最も安定した良い推測性能が得られ、モデル2よりもわずかに良い推測性能が得られた。つまり、今回の結果では学習時には全てのSSIを含んで学習した群の方が低いSSIを除外して学習した群よりも推測精度が高く、また眼内疾患を除外せずに学習した群は除外して学習した群よりもわずかに推測精度が高かった。

モデル間の精度を比較するための5分割交差検証の結果を受け、その推測性能を調べるため、今回の5分割交差検証にて比較的安定した結果が出たモデル1の条件、つまり学習時に眼内疾患及び全てのSSIを学習対象に含めた条件にて、10分割交差検証を用いて検証群とテスト群を分離したモデルを再度作成し直した。その主な評価結果を表5に示し、評価対象から合併疾患を除外した図を図2から図9に示す。その中の図2から図9は、評価対象のSSIを7以上に限定した図である。図10から図17はSSIの解析の図であるため、全てのSSIを評価対象にした。

条件	眼内疾患	SSI	評価項目	HFA24-2	HFA10-2
視野の閾値	眼内疾患 除外	SSI7 以上	測定点毎のピアソンの 積率相関係数	r=0.848 p<0.001	r=0.882 p<0.001
			測定点毎のスピアマンの順位相 関係数	ρ =0.814 p<0.001	ρ =0.781 p<0.001
			RMSE(dB)	3.27 ±2.20	3.02 ±2.28
			測定点毎の MAE(dB)	2.46 ±1.73	2.24 ±1.66
			視野全体の平均の MAE(dB)	1.62 ±1.70	1.39 ±1.46
			RMSE(dB)	3.81 ±2.43	3.72 ±2.72
	SSI7 未 満	測定点毎の MAE(dB)	2.96 ±2.03	2.85 ±2.13	
		視野全体の平均の MAE(dB)	2.01 ±2.02	1.86 ±1.94	
		RMSE(dB)	3.40 ±2.28	3.13 ±2.37	
	眼内疾患 含有	SSI7 以上	測定点毎の MAE(dB)	2.55 ±1.79	2.32 ±1.74
			視野全体の平均の MAE(dB)	1.68 ±1.76	1.46 ±1.56
			ピアソンの 積率相関係数	r=0.903 p<0.001	r=0.923 p<0.001
MD 値	眼内疾患 除外	SSI7 以上	スピアマンの 順位相関係数	ρ =0.804 p<0.001	ρ =0.844 p<0.001
			MAE(dB)	1.62 ±1.74	1.44 ±1.55
			測定点毎のピアソンの 積率相関係数	r=0.431 p<0.001	r=0.438 p<0.001
視野の 閾値の 進行速度	眼内疾患 除外	SSI7 以上	測定点毎のスピアマンの順位相 関係数	ρ =0.521 p<0.001	ρ =0.525 p<0.001
			測定点毎の MAE(dB)	0.594 ±0.738	0.819 ±0.900
			ピアソンの 積率相関係数	r=0.429 p<0.001	r=0.383 p<0.001
MD 値の進 行速度	眼内疾患 除外	SSI7 以上	スピアマンの 順位相関係数	ρ =0.496 p<0.001	ρ =0.481 p<0.001
			MAE(dB)	0.257 ±0.302	0.314 ±0.311
			測定点毎のピアソンの 積率相関係数	r=0.429 p<0.001	r=0.383 p<0.001

表5 10分割交差検証の評価結果。視野の閾値及び MD 値に関し、極めて高い推測性能が得られた。HFA24-2は最終視野が Anderson-Patella の基準に該当する眼のみを評価対象とした。

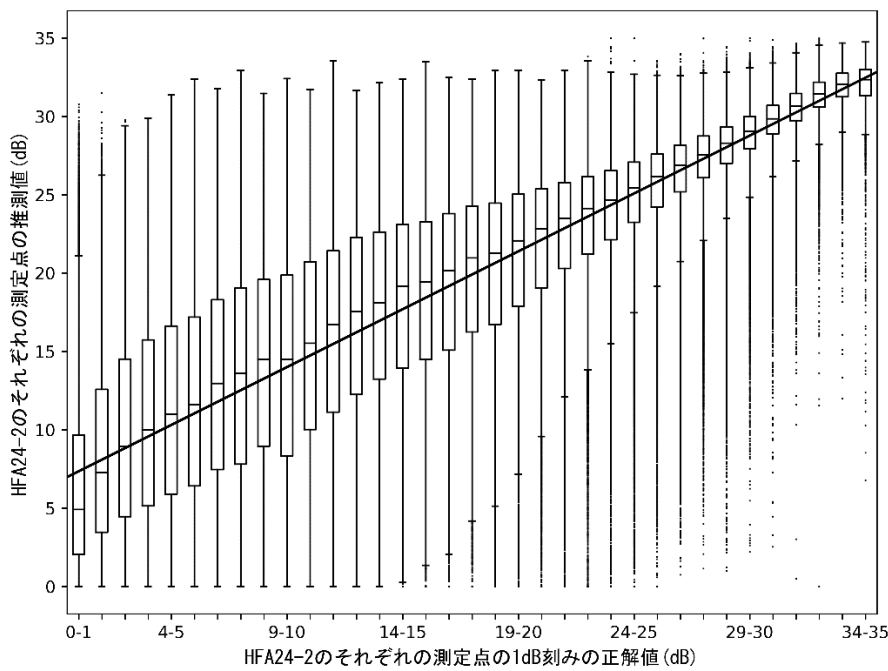


図2 OCTからのHFA24-2視野の測定点毎の推測値と正解値のグラフ。
高い相関が得られた。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.814$ 、
ピアソンの積率相関係数 $p < 0.001$, $r = 0.848$)

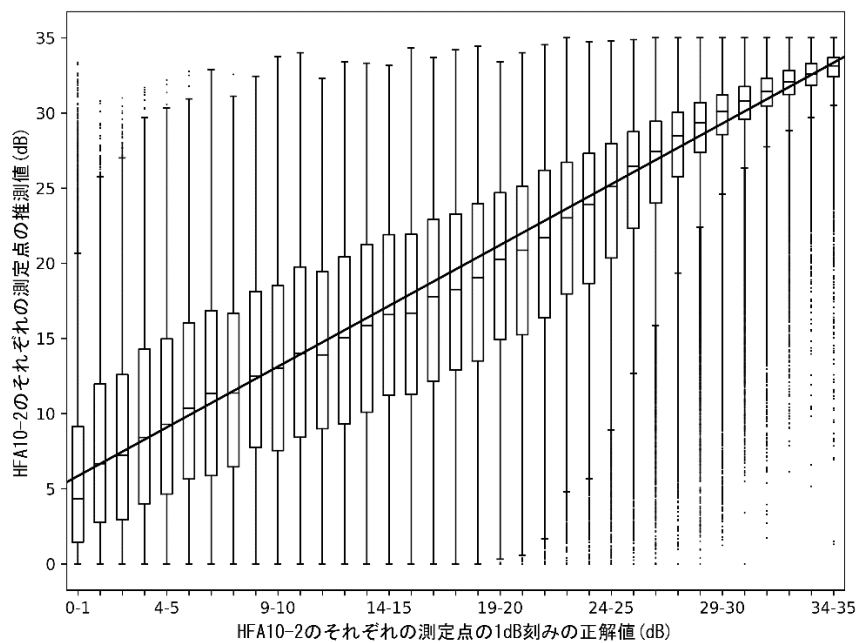


図3 OCTからのHFA10-2視野の測定点毎の推測値と正解値のグラフ。
HFA24-2同様、高い相関が得られた。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.781$ 、
ピアソンの積率相関係数 $p < 0.001$, $r = 0.882$)

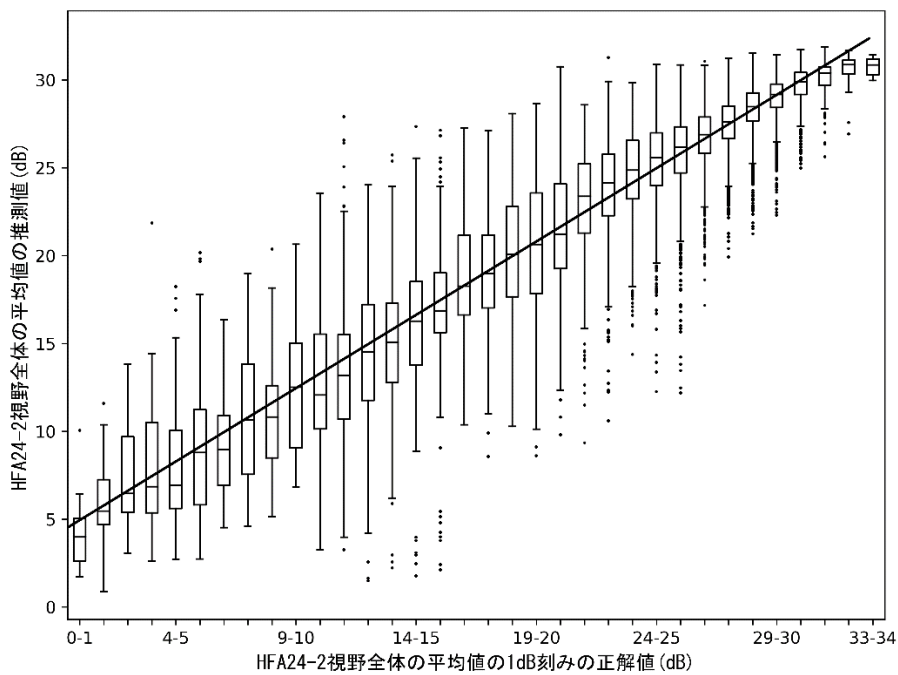


図4 OCTからのHFA24-2平均視野の推測値と正解値のグラフ。

極めて高い相関が得られた。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.844$ 、

ピアソンの積率相関係数 $p < 0.001$, $r = 0.899$)

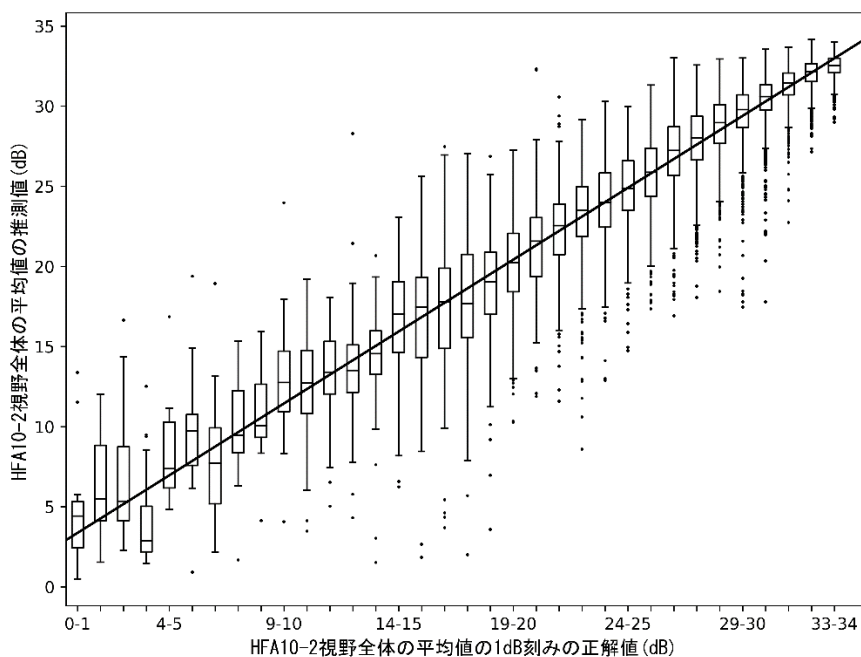


図5 OCTからのHFA10-2平均視野の推測値と正解値のグラフ。

HFA24-2同様、極めて高い相関が得られた。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.890$ 、

ピアソンの積率相関係数 $p < 0.001$, $r = 0.930$)

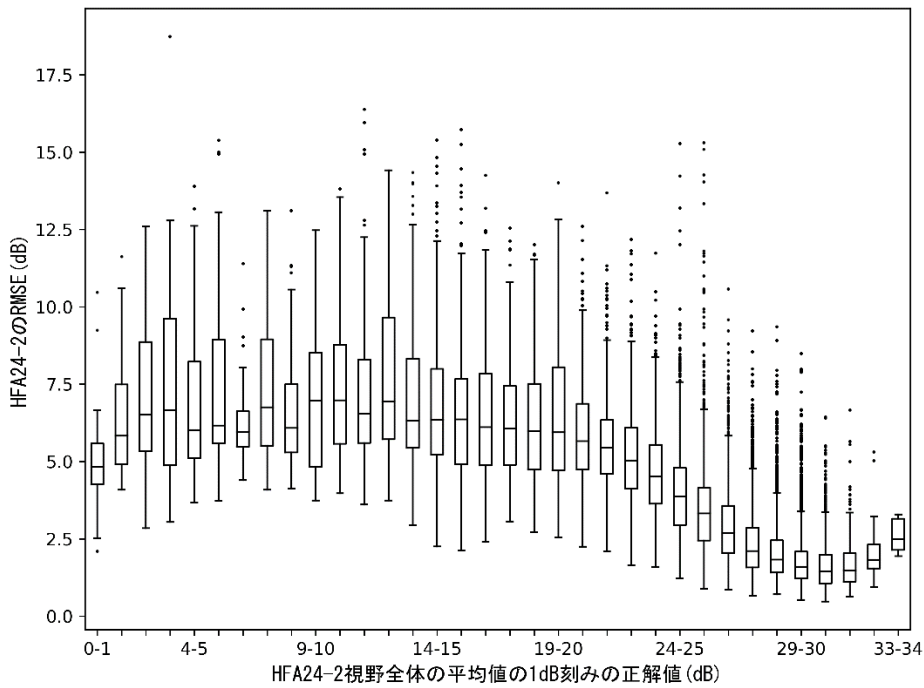


図6 HFA24-2視野全体の正解値と RMSE のグラフ。

平均視野が悪いほど推測が困難であった。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = -0.763$)

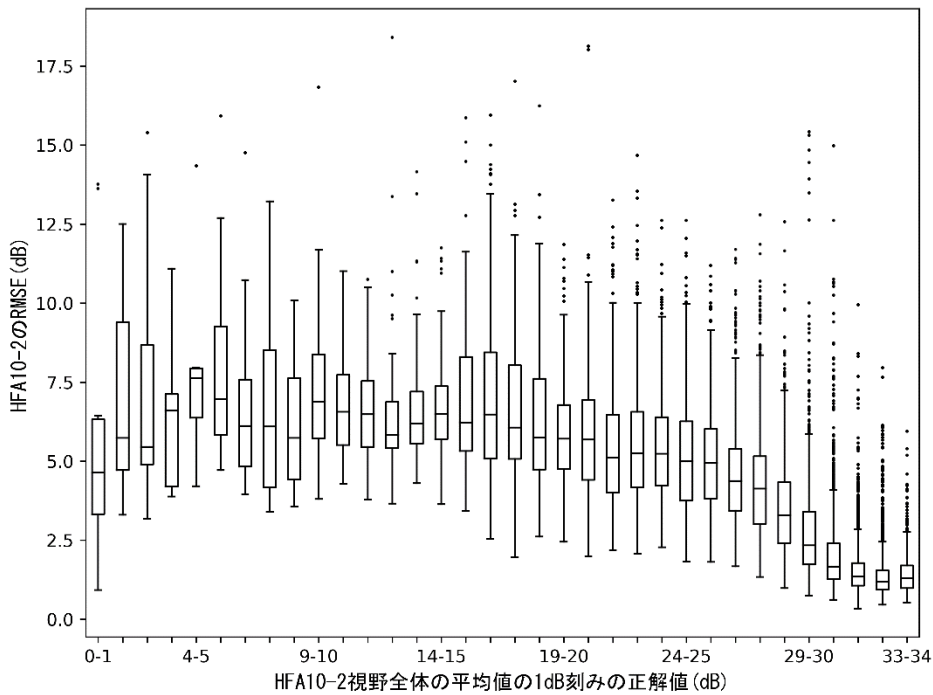


図7 HFA10-2視野全体の正解値と RMSE のグラフ。

HFA24-2同様、平均視野が悪いほど推測が困難であった。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = -0.797$)

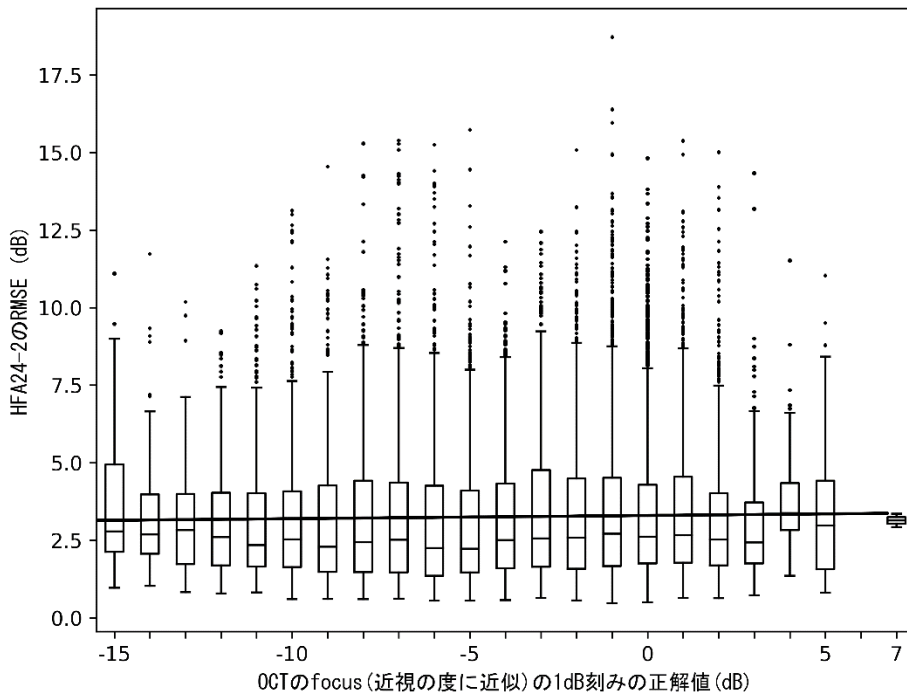


図8 OCTのfocus（近視の度に近似）とHFA24-2のRMSEのグラフ。
近視が強いほど高い推測精度が得られた。

（スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.0430$ ）

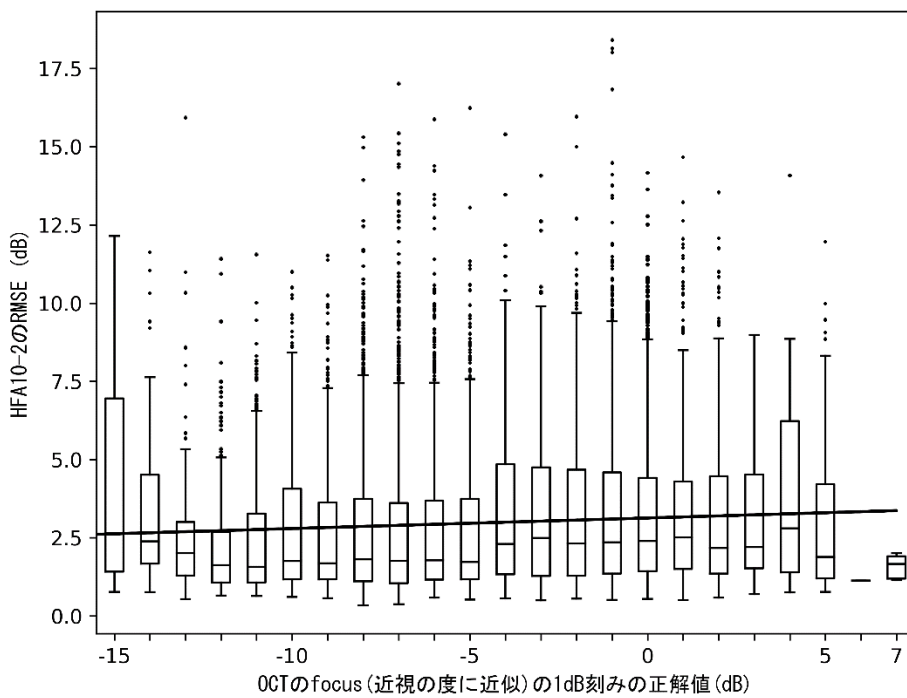


図9 OCTのfocusとHFA10-2のRMSEのグラフ。

HFA24-2同様、近視が強いほど高い推測精度が得られた。

（スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.102$ ）

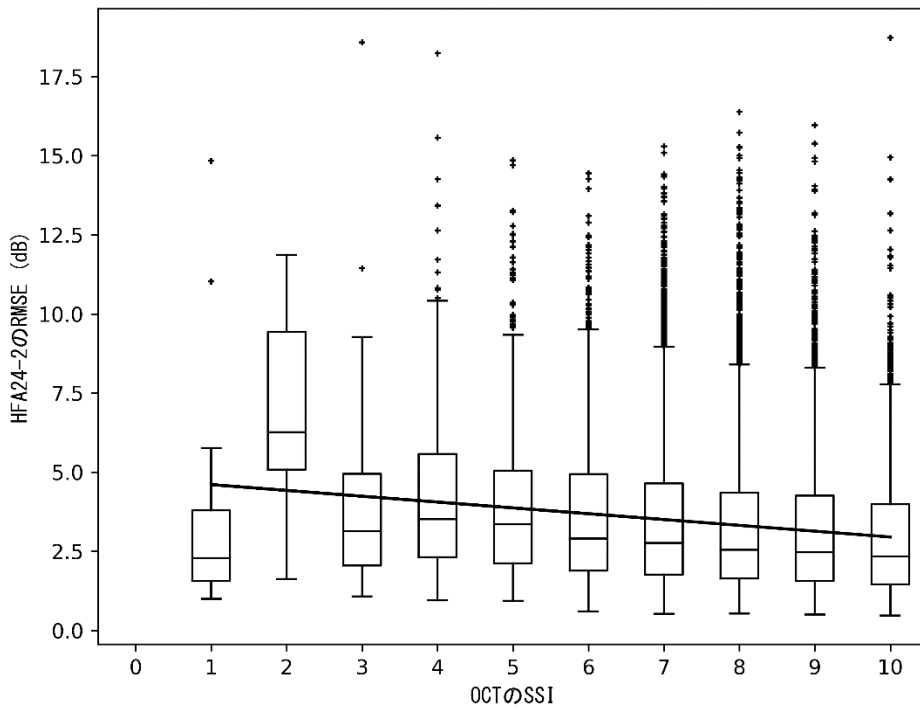


図10 OCTの全てのSSI（信号強度指数）とHFA24-2のRMSEのグラフ。

SSIが低い（＝ノイズが乗ったOCT）ほど推測が困難であった。

（スピアマンの順位相関係数 $p < 0.001$, $\rho = -0.115$ ）

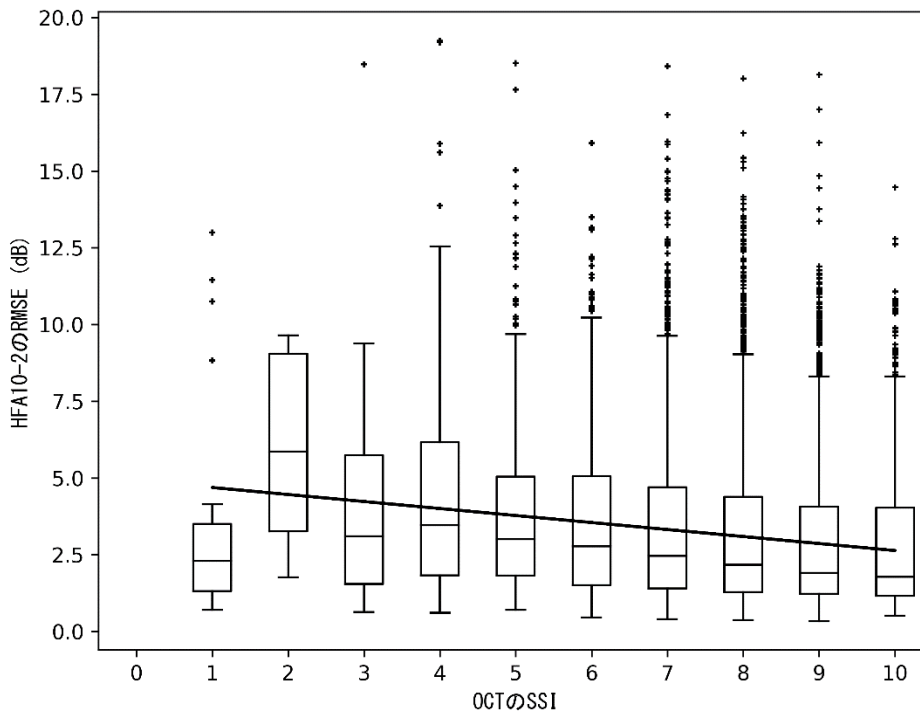


図11 OCTの全てのSSIとHFA10-2のRMSEのグラフ。

HFA24-2同様、SSIが低いほど推測が困難であった。

（スピアマンの順位相関係数 $p < 0.001$, $\rho = -0.128$ ）

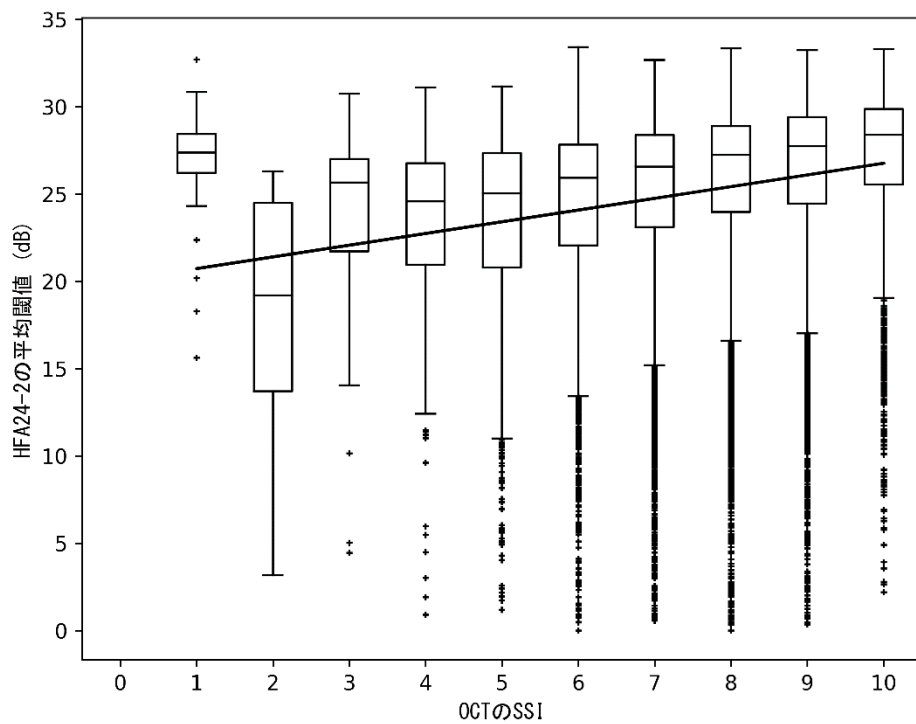


図12 OCTの全てのSSIとHFA24-2の平均閾値のグラフ。

SSIが低いほど平均閾値が低値であった。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.220$)

図の線は回帰直線。回帰係数 = 0.671 (dB/SSI)。

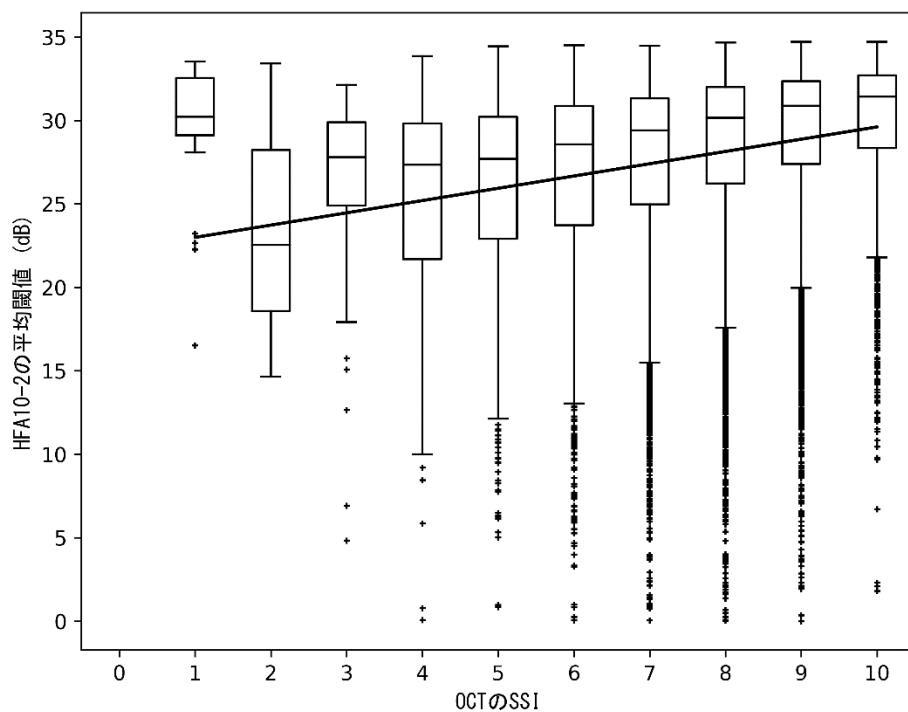


図13 OCTの全てのSSIとHFA10-2の平均閾値のグラフ。

HFA24-2同様、SSIが低いほど平均閾値が低値であった。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.230$)

図の線は回帰直線。回帰係数 = 0.736 (dB/SSI)。

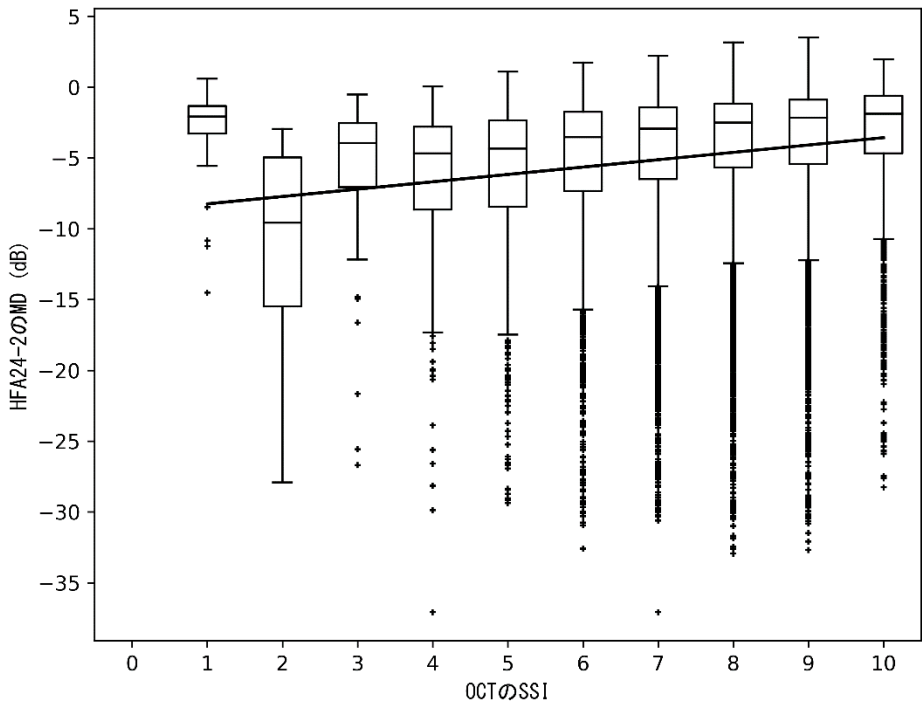


図14 OCTの全てのSSIとHFA24-2のMD値のグラフ。

SSIが低いほどMD値が低値であった。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.166$)

図の線は回帰直線。回帰係数 = 0.520 (dB/SSI)。

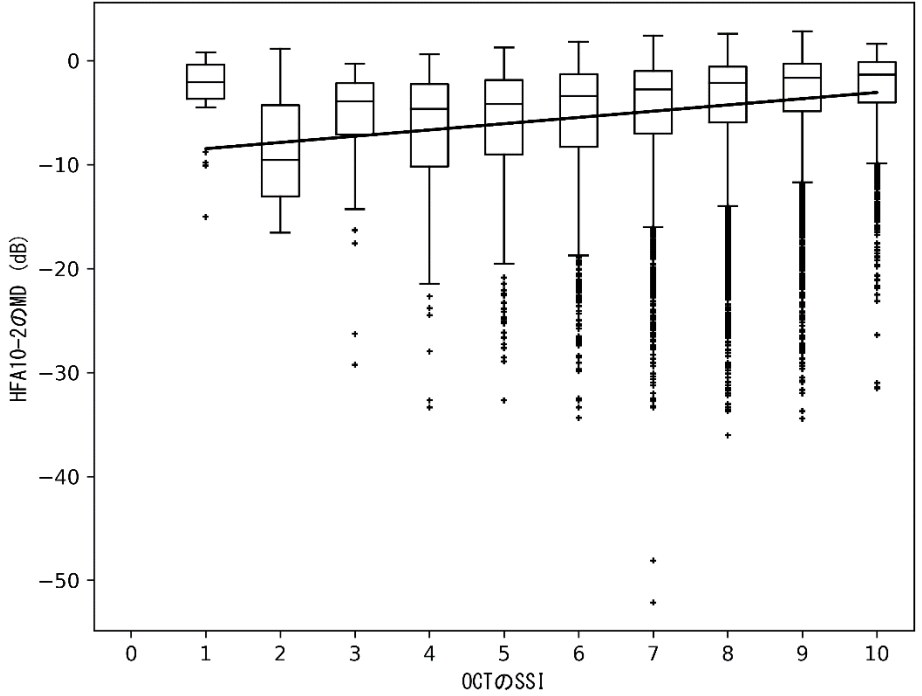


図15 OCTの全てのSSIとHFA10-2のMD値のグラフ。

HFA24-2同様、SSIが低いほどMD値が低値であった。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = 0.188$)

図の線は回帰直線。回帰係数 = 0.601 (dB/SSI)。

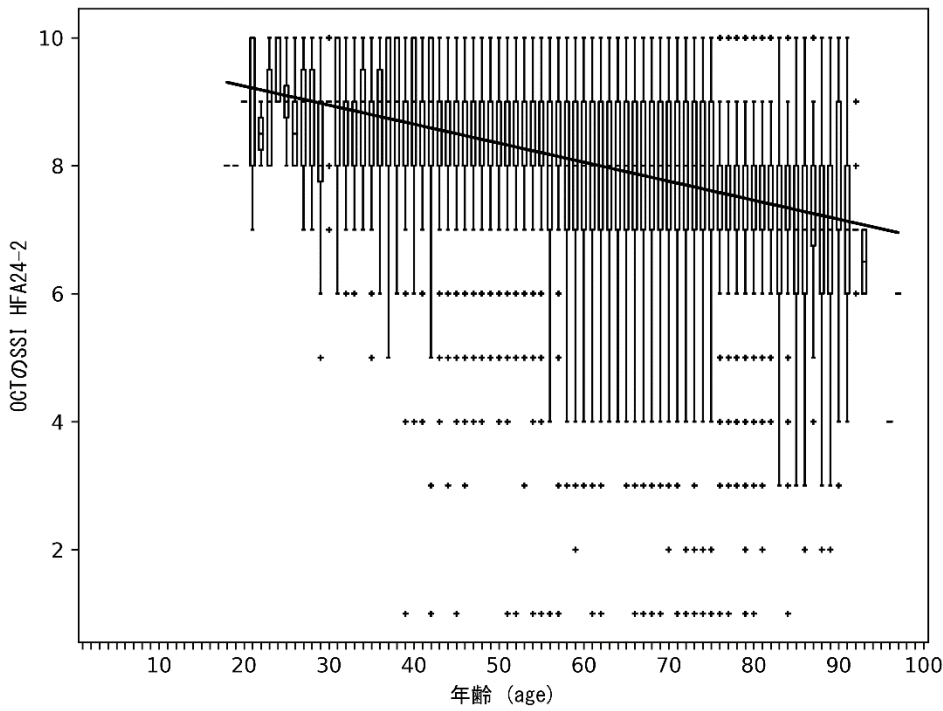


図16 HFA24-2のペアデータからの、年齢と全ての OCT の SSI の平均値のグラフ。

年齢が高齢になるほど SSI は低下していく傾向が認められた。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = -0.279$)

図の線は回帰直線。回帰係数 = -0.0297 (SSI/age)。

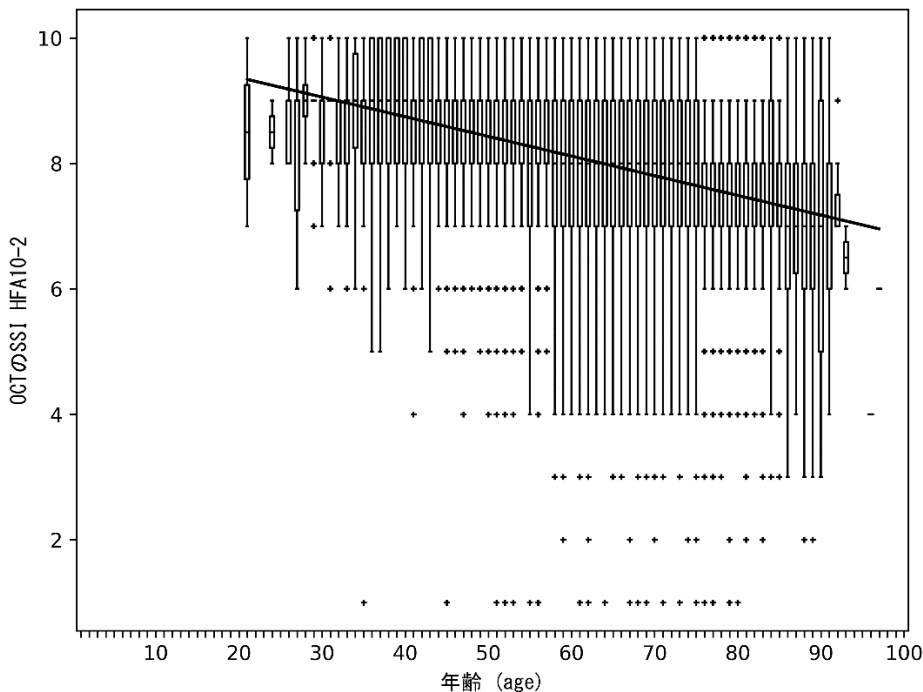


図17 HFA10-2のペアデータからの、年齢と全ての OCT の SSI の平均値のグラフ。

HFA24-2同様、年齢が高齢になるほど SSI は低下していく傾向が認められた。

(スピアマンの順位相関係数 $p < 0.001$, $\rho = -0.293$)

図の線は回帰直線。回帰係数 = -0.0314 (SSI/age)。

考察

OCT は視野検査よりも短い時間で測定できるため、OCT から視野を推測する事は臨床上重要である。より多くのペアデータを確保し、高い推測精度を得るため、従来の2次元のモデルの代わりに3DCNN を用いた。教師値は、測定毎のばらつきを抑えるために HFA の測定点毎の回帰直線から計算して用いた。正の傾きを0とした事による影響及び、SSIの低下による見かけ上の視野の進行の影響を減らすため、傾きの教師値に予め補正を行った。学習対象を変更する事によって得られる推測性能の違いを比較するため、それぞれの学習条件での5分割交差検証モデルを作成し、共通した評価対象での推測性能を比較した結果、今回は全ての眼内疾患を含有して全ての SSI を対象として学習した群が最も良い結果が得られた。5分割交差検証で最も良い結果が得られた学習条件と同一の学習条件で10分割交差検証モデルを作成し学習し、最終的な推測精度を計算し、モデルの構成要素である、focus、SSI、平均閾値と RMSE との関係を解析した。

今回のモデルによって、単回の OCT の黄斑マップから極めて精度の高い HFA24-2及び10-2の閾値推測が可能であった。OCT の黄斑マップ単独では範囲が狭いため、従来は HFA24-2の閾値推測は困難と考えられていた^{3,4,9}が、今回のモデルでは HFA24-2も高い推測性能が得られた。

学習対象を変化させて5分割交差検証を利用して学習した、それぞれの結果間に関して検定を行ってみると、RMSE のわずかな差でも多くの場合に有意差が得られたが、再度同様のモデルを作成して学習を行って検定を行うと逆の結果が出る事も頻繁にあった。その理由は、検定は「結果」に対するもので、「モデル」への検定では無いためである。モデルの優劣を厳密に検定するためには、同一のモデルを複数回学習し直す必要があるため、少なくとも合計で数ヶ月間以上の長い学習時間がかかると予想されるため、今回は行っていない。今回の結果からは、必ずしも眼内疾患を除外して学習する必要性は無いという可能性が示唆され、また、低い SSI を敢えて除外して学習する必要性は無いという可能性も示唆された。もし眼内疾患を除外する必要が無ければ、後ろ向きに眼内疾患の有無を手作業で調べる労力が減り、新たな症例を集めやすくなる可能性が考えられる。

過去に OCT から HFA の閾値を推測する報告は多数あり、そのほとんどは2次元のモデルであった。直近の既報の値は、HF10-2の RMSE:5.92dB (Xu ら)⁹に対して今回のモデルでは3.02dB (小さい方が良い)、HFA10-2 測定点毎の MAE:5.0dB (Asaoka ら)⁴に対して今回のモデルでは2.24dB (小さい方が良い)、HFA10-2視野全体の平均の MAE:2.6dB (Hashimoto ら、但し HFA24-2の値を与えた後では1.9dB)³に対して今回のモデルでは HFA24-2の値を与えずに1.39dB (小さい方が良い)、HFA10-2の測定点毎のピアソンの $r=0.61$ (Cirafici ら)⁵ に対して今回のモデルでは $r=0.882$ (大きい方が良い)、HF24-2の RMSE:4.51dB (Shin ら)⁶に対して今回のモデルでは3.27dB (小さい方が良い) であり、今回のモデルは既報の値をいずれも大きく上回る推測性能を達成した。その原因として主に考えられるのは、3DCNN を用いた事、ペアデータ数

が多い事、教師値に回帰直線を用いた事である。但し、既報とは母集団が異なり、既報では大学病院の重症例が多いため、より視野の推測が困難であったと考えられるために単純な比較はできない。

今回のモデルでは従来の一般的な考え²¹とは逆に、近視が強いほど高い推測精度が得られた。その主な理由として考えられるのは、3DCNN モデルを採用した事である。強度近視眼では網膜厚が薄い事が知られており²²、眼球の形状によって部位別に生理的な網膜厚が異なる可能性が考えられる。2次元のモデルでは、3次元の情報を2次元に圧縮する際に形状の情報の多くが失われるが、3DCNN モデルであれば、3次元の情報をそのまま学習する事ができるため、強度近視の症例により正確に推測できるようになったのではと考えた。

今回のモデルにより現在の視野だけではなく、同時に視野の進行速度も単回の OCT 画像からある程度推測できたが、単一の診療所の調査のため有意な進行速度を教師値に採用できた症例数が少ないため、推測性能は高く無い。既報では、既存の多数の視野の進行速度を教師値として視野からその進行速度を推測するという報告²³はあるが、OCT 画像を元にして視野の進行速度を推測するという報告は調べた限りでは見当たらない。視野のデータよりも OCT 画像のデータの方が情報量は多く、機械学習によって3次元の眼底構造から視野が進行しやすいかどうかを学習できる可能性が考えられ、さらに OCT 画像から視野を推測してその推測視野から既報と同様に視野の進行速度を推測する事も可能であるため、今後多数の症例が集まれば、視野を元にするよりも OCT 画像を元にしたほうが、視野の進行速度の推測精度の向上が見込める可能性も考えられる。また、今回の研究では単回の OCT 画像を元に進行速度を推測しており、OCT 画像の時系列データによる情報は加味していない。将来的には、今回の研究モデルを基本モデルとし、時系列での OCT 画像データを今回のモデルに入力して、出力された時系列でのデータに対して追加解析を行う新たな機械学習モデルを作成する事により、進行速度の推測精度をさらに向上できる可能性も考えられる。

当院では既に今回のモデルを臨床応用しており、OCT 撮影後の画像を出力した直後に自動的に 24-2、10-2の視野の推測 + 視野の進行速度の推測を行っているが、他院での使用のためには、OCT の機種や撮影条件が少しでも異なると正しく推測できないため、多くのメーカーの OCT で様々な撮影条件での学習が必要である。そのためには多施設共同研究が必須であり、今回の 3DCNN による後ろ向き全例調査は、そのような将来の多施設共同研究の基礎となる手法として有用な可能性が考えられる。

利益相反：該当無し

参考文献

1. Weinreb, R. N. *et al.* Primary open-angle glaucoma. *Nature Reviews Disease Primers* vol. 2 Preprint at <https://doi.org/10.1038/nrdp.2016.67> (2016).
2. Altangerel, U., Spaeth, G. L. & Rhee, D. J. *Visual function, disability, and psychological impact of glaucoma. Curr Opin Ophthalmol* vol. 14 (2003).
3. Hashimoto, Y. *et al.* Predicting 10-2 Visual Field From Optical Coherence Tomography in Glaucoma Using Deep Learning Corrected With 24-2/30-2 Visual Field. *Transl Vis Sci Technol* **10**, (2021).
4. Asaoka, R. *et al.* A Joint Multitask Learning Model for Cross-sectional and Longitudinal Predictions of Visual Field Using OCT. *Ophthalmology Science* **1**, 100055 (2021).
5. Cirafici, P. *et al.* Point-wise correlations between 10-2 Humphrey visual field and OCT data in open angle glaucoma. *Eye (Basingstoke)* **35**, 868–876 (2021).
6. Shin, J., Kim, S., Kim, J. & Park, K. Visual field inference from optical coherence tomography using deep learning algorithms: A comparison between devices. *Transl Vis Sci Technol* **10**, (2021).
7. Sugiura, H. *et al.* Estimating glaucomatous visual sensitivity from retinal thickness with pattern-based regularization and visualization. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 783–792 (Association for Computing Machinery, 2018). doi:10.1145/3219819.3219866.
8. Xu, L. *et al.* Improving Visual Field Trend Analysis with OCT and Deeply Regularized Latent-Space Linear Regression. *Ophthalmol Glaucoma* **4**, (2021).
9. Xu, L. *et al.* PAMI: A Computational Module for Joint Estimation and Progression Prediction of Glaucoma. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 3826–3834 (Association for Computing Machinery, 2021). doi:10.1145/3447548.3467195.
10. Xu, L. *et al.* Predicting the Glaucomatous Central 10-Degree Visual Field From Optical Coherence Tomography Using Deep Learning and Tensor Regression. *Am J Ophthalmol* **218**, (2020).
11. Hashimoto, Y. *et al.* Deep learning model to predict visual field in central 10° from optical coherence tomography measurement in glaucoma. *British Journal of Ophthalmology* **105**, (2021).
12. Asano, S. *et al.* Predicting the central 10 degrees visual field in glaucoma by applying a deep learning algorithm to optical coherence tomography images. *Sci Rep* **11**, (2021).
13. Anderson DR, P. V. *Automated Static Perimetry, 2nd edition.* (Mosby, St. Louis, 1999).
14. Yu, S. *et al.* Comparison of SITA Faster 24-2C test times to legacy SITA tests. *Invest Ophthalmol Vis Sci* **60**, (2019).
15. L.J., S., R.A., R. & D.P., C. Standard or Fast?-Differences in precision between SITA Standard and SITA Fast testing algorithms and their utility for detecting visual field deterioration. *Invest Ophthalmol Vis Sci* **55**, (2014).
16. Saunders, L. J., Russell, R. A. & Crabb, D. P. Measurement precision in a series of visual fields acquired by the standard and fast versions of the swedish interactive thresholding algorithm analysis of large-scale data from clinics. *JAMA Ophthalmol* **133**, (2015).
17. Nakatani, Y., Higashide, T., Ohkubo, S., Takeda, H. & Sugiyama, K. Effect of cataract and its removal on ganglion cell complex thickness and peripapillary retinal nerve fiber layer thickness measurements by fourier-domain optical coherence tomography. *J Glaucoma* **22**, 447–455 (2013).
18. Kim, M., Eom, Y., Song, J. S. & Kim, H. M. Effect of Cataract Grade according to Wide-Field Fundus Images on Measurement of Macular Thickness in Cataract Patients. *Korean Journal of Ophthalmology* **32**, (2018).
19. shijianjian. EfficientNet-PyTorch-3D. github.com. Updated July 20, 2021. Accessed May 10, 2022. <https://github.com/shijianjian/EfficientNet-PyTorch-3D> (2021).
20. Tan, M. & Le, Q. v. EfficientNet: Rethinking Model Scaling for Convolutional Neural

Networks. *arXiv preprint:1905.11946* (2019).

21. Akashi, A. *et al.* The ability of SD-OCT to differentiate early glaucoma with high myopia from highly myopic controls and nonhighly myopic controls. *Invest Ophthalmol Vis Sci* **56**, (2015).
22. Sezgin Akcay, B. I., Gunay, B. O., Kardes, E., Unlu, C. & Ergin, A. Evaluation of the Ganglion Cell Complex and Retinal Nerve Fiber Layer in Low, Moderate, and High Myopia: A Study by RTVue Spectral Domain Optical Coherence Tomography. *Semin Ophthalmol* **32**, (2017).
23. Murata, H., Araie, M. & Asaoka, R. A new approach to measure visual field progression in glaucoma patients using variational bayes linear regression. *Invest Ophthalmol Vis Sci* **55**, (2014).