

論文の内容の要旨

論文題目 MRI 画像を用いた radiomics と機械学習による
神経膠腫の悪性度及び *IDH* 変異の予測

氏名 高橋 慧

要旨

序文:

神経膠腫 (glioma) は原発性脳腫瘍の中で髄膜腫と並び最も高頻度に見られる。神経膠腫は heterogeneous な腫瘍の集合であり、組織像を横軸とし、悪性度 (grade) を縦軸として分類される。悪性度が最も高い膠芽腫 glioblastoma (GBM) は、新しい治療方法が次々に提案されているものの、5 年生存率は未だに 10%程と非常に悪い。また、分子遺伝学的には、神経膠腫において *IDH1/2* 変異、*ATRX* 変異、*TP53* 変異、染色体 1p/19q 共欠失などが重要な遺伝子異常として知られる。*IDH1* 変異およびそのホモログである *IDH2* 変異はその中でも最も重要で最も根幹となる遺伝子異常である。*IDH1* 変異は lower grade glioma (LrGG) の 50-80%、GBM の 12%に認められ、予後に強く相関している。

神経膠腫の診断のゴールドスタンダードは病理標本によるものであるが、外科的侵襲を伴うことや、原理的に腫瘍の一部分のみの診断になることなどの問題点がある。さらには悪性度や *IDH* 変異の有無は、手術戦略を含め治療方針を決定づけ得るので、それらを治療介入前に予測できることが望ましい。

MRI は術前診断、治療計画、臨床管理、治療判定等に日常的に使用されている。MRI は非侵襲的であることが大きなメリットである。

非正規分布拡散に基づいた MRI 撮像方法に拡散クルトシス画像 (diffusion kurtosis imaging, DKI) がある。DKI は拡散テンソル画像 (diffusion tensor imaging, DTI) よりも、理論上は正確に水分子の拡散の状態を表現されているとされ、DKI のいくつかのパラメータは神経膠腫の悪性度診断に有用であるとの報告がある。

しかし、DKI を用いた神経膠腫の悪性度診断に関する既存の研究は、画像より抽出したいくつかのパラメータについてのみ検討したものである。これは、多数の特徴量を網羅的に探索し、検討するアプローチではない。このようなアプローチ方法は radiomics と呼ばれている。Radiomics は医療画像を絵としてではなく、データとして扱う、新たな学術分野を指し示すと同時に、医療用画像から多量のデータを抽出、分析し何らかの予測を行う workflow を表現する言葉である。

我々の研究の目的は、 $T2b_0$ 画像 ($T2b_0$)、DWI,DTI, DKI より radiomics の workflow を用い抽出した多数の特徴量を対象に機械学習を行い、神経膠腫の悪性度診断及び *IDH* 変異予測を行う診断モデルを作成することである。

方法:

東京大学医学部附属病院にて 2014 年 9 月から 2018 年 1 月までの期間に摘出術または生検術が行われ、病理診断にて神経膠腫と診断が確定した症例のうち、MRI 画像が以下の 2 つの条件を満たしている症例を研究対象とした。

1. 術前に T2 ρ , DWI, apparent diffusion coefficient (ADC, DTI から得られるパラメータの 1 種), fractional anisotropy (FA, 同じく DTI から得られるパラメータの 1 種), そして、mean kurtosis (DKI から得られるパラメータの 1 種。以下 MK と称する。) の全てのシーケンスが撮像されている。

2. MRI Gadolinium-enhanced T1WI (GdT1) と fluid attenuated IR (FLAIR) が両方ともまたはいずれかが撮像されている。

また、2014 年 9 月から 2017 年 7 月までの症例を学習用のデータセットに、2017 年 8 月から 2018 年 1 月までの症例を独立テスト用のデータセットとした。

悪性度に関しては、機械学習に足る症例数を担保するために、grade IV の GBM と grade III と grade II の LrGG の二群に分けて解析を行うこととした。IDH 変異はすべて Sanger sequence にて決定した。

Volume of interest (VOI) 付きの画像を使用し、1 種類の撮影方法あたり 8+9×(10+11+13+13+5) = 476 種類、1 症例あたり 476×6 = 2856 種類の特徴量を抽出することとした。

結果:

悪性度診断

55 例が対象となり、内訳は 14 症例の grade II の神経膠腫 (びまん性星細胞腫 diffuse astrocytoma 3 例、乏突起膠腫 oligodendroglioma 11 例)、12 症例の grade III の神経膠腫 (退形成性星細胞腫 anaplastic astrocytoma 7 例、退形成性乏突起膠腫 anaplastic oligodendroglioma 5 例)、grade IV の症例が 29 例 (全て GBM) であった。磁場強度 1.5T の条件で 27 症例、磁場強度 3.0T 条件で 28 症例が撮像されていた。1 症例のみ (diffuse astrocytoma) 1.5T、3.0T の 2 つの条件で撮像されていた。独立テスト用のデータセットは 11 例であり、全て 3.0T で撮像されていた。

最も精度の高い悪性度診断の機械学習モデルは、 $k_{ROIGSumAverage}$, $k_{LLLLGRE}$, k_{HHLZSN} , $k_{LHLStrength}$, $d_{LHLEntropy}$, $d_{LLLGCORRELATION}$ の 6 つの特徴量を使い SVM (kernel is rbf, c is 1.0) を使用した場合で、独立テスト用データセットに対する予測率は 0.91 (10/11)、AUC は 0.93 ± 0.03 であった

IDH 変異予測

磁場強度 1.5T の条件で 26 例、3.0T の条件で 24 例の計 50 例が対象となった。1 症例のみ (grade II diffuse astrocytoma、悪性度診断と同様の症例) 1.5T、3.0T の 2 つの条件

で撮像されていた。内訳は *IDH* 野生型が 25 例 (anaplastic astrocytoma 2 例、glioblastomas 23 例)、*IDH1* 変異型が 25 例 (diffuse astrocytoma 3 例、anaplastic astrocytoma 4 例、oligodendroglioma 10 例、anaplastic oligodendroglioma 4 例、glioblastoma 4 例) であった。*IDH* 変異は全て *IDH1* 変異 (R132H) であり、*IDH2* 変異は認めなかった。悪性度診断と同様に独立テスト用のデータセットの 11 例は全て 3.0T で撮像されていた。

最も精度の高い *IDH* 変異予測モデルは *k_LHLStrength*, *b2_LLLGVariance*, *b2_HLLGCorrelation*, *b2_LLLLRHGE* の 4 つの特徴量を使用し、SVM (kernel は sigmoid, *c* は 500, *gamma* は 0.0001) を用いた場合で、独立テスト用データセットに対する予測精度は 0.82 (9/11) であり、AUC は 0.92 ± 0.03 であった。

考察：

本研究では、MRI 画像から抽出した特徴量のみを用いて高精度に、神経膠腫の悪性度診断及び *IDH* 変異予測を行う機械学習モデルを作成した。特に有用な特徴量は悪性度診断では ADC と MK から、*IDH* 変異予測では *b2000DWI* と MK からそれぞれ抽出された。

Radiomics とは、放射線画像を画像ではなくデータとして扱う新しい学術分野である。画像データから多数の特徴量を抽出し解析することで、臨床的、病理的な事象を予測する。特に画像データを入力として、遺伝子異常を予測する場合は radiogenomics と呼ばれる。本研究は、*T2b0*, DTI 及び DKI の多種類の MRI シークエンスを用いて神経膠腫の悪性度診断及び *IDH* 変異予測の機械学習モデルを作成し、その有用性を比較検討した、渉猟し得る限り初めての研究である。

高 *b* 値 DWI は水分子の slow components の描出により優れている可能性が示唆されており、高 *b* 値 DWI が *IDH* 変異予測に有用であると期待されている。

本研究においても、高 *b* 値 DWI から抽出された特徴量は *IDH* 変異予測に有用である可能性が示された。

ADC は拡散の大きさを、FA は拡散の異方性をそれぞれ表していると考えられている。ADC は細胞密度を反映しており ADC は神経膠腫の悪性度診断に有用であると言われている。Han らと Kang らは高い *b* 値の DWI より算出された ADC が grading に有用であることを示したが、これは我々の結果と矛盾しない。

MK は水分子の拡散がどの程度制限されているかを画像化したものになる。理論的には、DKI は DTI よりも生体内での水分子の挙動をより正確に反映する。DKI、特に MK が神経膠腫の悪性度診断に有用であることは既にいくつもの研究で示唆され、メタ解析によっても示されている。本研究でも MK より抽出された特徴量が悪性度診断を行う機械学習モデルを作成する際に重要な役割を果たしている。*IDH* 変異予測については、Bisdas らが *IDH* 変異の予測に DKI が有用である可能性を示したが、*IDH* 変異の有無は主に免疫染色によって決定され、独立した試験データセットによる検証は行われていない。我々は、サンガー法

により *IDH* 変異の有無を決定したデータセットから、MK より抽出された特徴量である *k_LHLstrength* が *IDH* 変異予測モデルの精度向上に寄与していることを見出し、さらに独立した試験データセットを用いて示した。

特徴量は *semantic* と *agnostic* なものに大別される。*Semantic* な特徴量とは位置や、形状、大きさ等、一見して人間に認知可能な特徴量である。一方、*agnostic* な特徴量は画像に解析処理を加えたものになり、画像を眺めるだけで特徴量の違いを認知することは難しい。*Agnostic* な特徴量はさらに一次および二次以上に分けられる。一次の *agnostic* な特徴量は、個々のボクセルの値の分布に対しヒストグラム解析を行うことで得られる。一方、二次以上の *agnostic* な特徴量の抽出方法で、よく用いられるものの一つにテクスチャ解析がある。本研究において、悪性度診断及び *IDH* 変異予測に特に有用である可能性が示された特徴量は、一次の *agnostic* な特徴量である *d_LHLEntropy* を除いて、全て二次以上の *agnostic* な特徴量であった。

本研究にはいくつかの課題が存在するものの、DWI, DTI, DKI の *radiomics* 解析と機械学習モデルの作成が神経膠腫の悪性度診断と *IDH* 変異予測に有用である可能性が示された。