

令和 5 年 6 月 1 日現在

機関番号：17102
研究種目：基盤研究(C)（一般）
研究期間：2020～2022
課題番号：20K07637
研究課題名（和文）人工知能による深層機械学習システムを用いた乳癌組織画像解析アルゴリズムの構築

研究課題名（英文）Construction of Breast Cancer Tissue Image Analysis Algorithm Using Deep Machine Learning System with Artificial Intelligence

研究代表者
甲斐 昌也（KAI, Masaya）
九州大学・医学研究院・共同研究員

研究者番号：10755242
交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究では、Whole Slide Image (WSI) を用いてConvolutional Neural Network (CNN) により、深層学習を行うことで、生物学的特性（サブタイプ）規定因子としての「ER」「PgR」「HER2」「Ki67」を評価することとした。ERに関しては、本症例数でも精度69.46%を得ることができた。このことは、高い感度・特異度をもって、HE染色像から、ER発現の予測をすることができることが期待された。一方、PgR、Ki67はERと比較し低い結果であった。HER2については、細胞膜染色であり、核染色とはアルゴリズムが異なるため評価を行うことは困難であった。

研究成果の学術的意義や社会的意義

乳癌においては、NSGによる生物学的特性の解析が発展し、今後はその特性に基づいた個別化医療が進められることが予想される。すなわち、膨大な臨床病理学的データをもとに、治療方針を判断しなければならなくなり、従来の人の手や目を中心とした解析作業には限界があると考えられるため、データを、正確・高速に処理し、診断から治療につなげるシステムの構築が急務であり、診断の効率化と個別化治療開発は、同時進行として進められるべき命題であると考えた。本研究では、AIによるDeep learningを行うことで、形態からタンパク発現の予測が、一定の精度と効率化をもって、診断につなげることができるとことが示唆された。

研究成果の概要（英文）：In this study, by performing deep learning with a Convolutional Neural Network (CNN) using Whole Slide Image (WSI), ER, PgR, HER2, Ki67 was evaluated. We were evaluated 69.46% accuracy was obtained for ER, even with this number of cases. This indicates that the HE stained images can predict ER expression with high sensitivity and specificity. On the other hand, the accuracy of PgR and Ki67 was lower than that of ER. It was difficult to evaluate HER2 because it is a cell membrane staining and the algorithm is different from nuclear staining. In conclusion, the present study shows that some of the markers targeted as subtype determinants can be predicted from HE staining. For HER2, it was difficult to evaluate because it is a plasma membrane stain and the algorithm is different from that of nuclear staining such as ER, PgR, and Ki67. In conclusion, the present study shows that some of the markers targeted as subtype-specific factors can be predicted from HE staining.

研究分野：医歯薬学

キーワード：乳癌 人工知能 免疫染色 個別化医療

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

乳癌は、日本人女性において最も罹患率の高い癌腫である。標準治療の確立により予後の改善は得られているが、それでも治療抵抗性を示す集団が存在し (Waks, JAMA. 2019) より細分化した個別化治療が求められている。乳癌治療においては正確な病理診断による生物学的特性の評価とサブタイプに基づいて治療方針が決定されるが、標的分子の多様性など複雑化の一途を辿っており、人間による診断には一定の限界があると考えられる。AI は、知的行動を人間に代わってコンピュータに行わせる技術である (Wolfgang, 1985)。医療の分野でも、AI を用いて放射線や内視鏡による画像診断や病理組織診断を行う試みが多くなされている

(Campanella Nat Med. 2019)。AI による診断支援技術の進歩は目覚ましく、その実用化に伴い、乳癌診療においても診断から治療への流れが大きく変化する可能性がある。次世代シーケンス (Next Generation Sequencing; NGS) は、膨大なシーケンシング反応を同時並行して実行できる技術であり、がん遺伝子パネル検査が乳癌の診断と治療に不可欠になっていくものと思われる (Crawford, Breast Cancer Res. Treat. 2017)。このような背景から、臨床病理ではタンパクから遺伝子レベルに至るまでの膨大なデータを、正確・高速に処理し、診断から治療につなげるシステムの構築が急務であり、診断の効率化と個別化治療開発は、同時進行として進められるべき命題である。またこれらは、病理医の負担を軽減し、人間と AI の利点・欠点を相補するものでもありうる。

2. 研究の目的

「AI により、乳癌組織 HE 染色画像から遺伝子・タンパク発現解析は可能である」という仮説の証明を行う。具体的には以下の 3 点に焦点を当てる。

1. 深層機械学習システムを用いた乳癌組織 HE 染色画像解析のアルゴリズムを構築する。
2. 治療標的となりうる多遺伝子・タンパクを特定する。
3. 個別化治療を目指した AI による病理画像解析の新たな活用への道筋をたてる。

3. 研究の方法

乳癌におけるデジタル病理画像解析の精度検証およびデータベースを構築する。すでに構築されたアルゴリズムを応用し、学習は教師データによる蓄積を行い、評価はテストデータとして Whole slide imaging (WSI) として電子化されている 1000 例以上の生検・手術検体のデジタル HE 染色画像を読み込ませ、精度を検証する。既知の臨床情報と対比させたデータベースを構築する。

サブタイプ規定因子における画像解析精度の検証を行う。乳癌のサブタイプを決定するうえで重要な因子であるエストロゲン受容体 (ER)、プロゲステロン受容体 (PgR)、HER2/neu タンパク、Ki67 のデジタルデータ化とその読み込みを行い、で学習させた H&E 染色画像と対比しさらに学習させる。

解析のターゲットをさらに拡大し、それぞれの解析精度の検証を行う。我々はこれまで、難治性乳癌に関係する因子として、形態形成シグナル (Hedgehog, Notch) や、乳癌幹細胞因子 (CD24, CD44)、乳癌遺伝子変異 (BRACAness)、腫瘍免疫関連因子 (TIL, PD-L1, Tbet, Granzyme B) に注目し、その意義について報告してきた。この深層学習システムを応用することで、難治性乳癌に関連するこれらの因子についても同様に解析し、その有用性の幅の検証を行う。

解析結果とがんゲノム情報を照らし合わせ、その特徴を検討する。我々はすでに乳癌組織における次世代シーケンシング (NGS) を用いた遺伝子パネルのデータベースを構築している。機械学習させた H&E 染色からの情報をもとに、NGS との対比を行い、その精度について検証する。

4. 研究成果

(1) デジタル病理画像解析

Whole Slide Image (WSI) を用いて、Convolutional Neural Network (CNN) により、深層学習を行った。HE 染色の WSI 上で癌部をマスキングし、学習させた。

(2) 生物学的特性の予測

生物学的特性 (サブタイプ) 規定因子としての「ER」「PgR」「HER2」「Ki67」を評価することとし

た。当初の目標症例数は「100例」を予定していたが、約半数にあたる43例で行った。ERに関しては、本症例数でも精度69.46%を得ることができた。このことは、高い感度・特異度をもって、HE染色像から、ER発現の予測をすることができることが期待された。一方、PgR, Ki67はそれぞれ精度48.89%、51.58%とERと比較し低い結果であった。これは今後、解析症例数を増やすことで精度を上げることができると考えている。HER2については、細胞膜染色であり、核染色のER, PgR, Ki67とはアルゴリズムが異なるため、現時点では評価はまだ行えていない。以上より、今回、サブタイプ規定因子として対象としたマーカーとしては、一部のタンパク発現についてはHE染色から予測が可能となる道筋が示された。

これらの結果は、以下の学術集会で発表を行った。

第121回日本外科学会定期学術集会

2021年4月10日

Deep Learningを用いた乳癌HE染色画像解析とタンパク発現予測に関する研究

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計3件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 原田由利菜, 久保真, 甲斐昌也, 山田舞, 森瞳美, 倉田加奈子, 川地眸, 金城和寿, 林早織, 島崎亜希子, 森崎隆史, 小田義直, 中津川宗秀, 中村雅史,
2. 発表標題 Deep Learningを用いた乳癌HE染色画像解析とタンパク発現予測に関する研究
3. 学会等名 第28回日本乳癌学会学術総会
4. 発表年 2020年

1. 発表者名 原田由利菜, 中津川宗秀, 久保真, 甲斐昌也, 山田舞, 森瞳美, 川地眸, 金城和寿, 林早織, 島崎亜希子, 森崎隆史, 岩崎健, 山元英崇, 小田義直, 中村雅史
2. 発表標題 Deep Learningを用いた乳癌HE染色画像解析
3. 学会等名 第57回九州外科学会
4. 発表年 2021年

1. 発表者名 原田由利菜, 中津川宗秀, 久保真, 甲斐昌也, 山田舞, 森瞳美, 川地眸, 金城和寿, 林早織, 島崎亜希子, 森崎隆史, 岩崎健, 山元英崇, 小田義直, 中村雅史
2. 発表標題 Deep Learning を用いた乳癌 HE 染色画像解析とタンパク発現予測に関する研究
3. 学会等名 第121回日本外科学会定期学術集会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	中村 雅史 (NAKAMURA Masafumi) (30372741)	九州大学・医学研究院・教授 (17102)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	久保 真 (KUBO Makoto) (60403961)	九州大学・医学研究院・准教授 (17102)	
研究分担者	小田 義直 (ODA Yoshinao) (70291515)	九州大学・医学研究院・教授 (17102)	
研究分担者	中津川 宗秀 (NAKATSUGAWA Munehide) (70448596)	東京医科大学・医学部・准教授 (32645)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関