

日本医療研究開発機構 医療機器開発推進研究事業
事後評価報告書



I 基本情報

研究開発課題名:

(日本語) LC-SCRUM レジストリを用いた形態学的バイオマーカーデジタルAI判定機器の治療効果予測性能検証試験

(英語) Retrospective performance study of digital AI-biomarker system for cancer treatment stratification using LC-SCRUM registry

研究開発実施期間: 平成4年12月7日～令和6年3月31日

研究開発代表者 氏名: (日本語) 元井紀子

(英語) Noriko Motoi

研究開発代表者 所属機関・部署・役職:

(日本語) 地方独立行政法人埼玉県立病院機構 埼玉県立がんセンター・病理診断科・副部長 兼 がんゲノム医療センター・センター長

(英語) Saitama Cancer Center, Department of Pathology, Deputy director and Center for Cancer Genomic Medicine, Director

II 研究開発の概要

和文

研究の背景

がん医療の最適化には、有効な治療法を正しく患者に届ける必要がある。本研究は、肺癌など難治固形がんにおける治療効果予測バイオマーカーの開発を通して、よりよいがん診療の実現を目指すものである。肺癌は、分子標的薬、免疫チェックポイント阻害剤 (ICI) など治療薬開発が進んでいるが、効果のある患者は限られている。特に、ICI は薬効予測が難しく、PD-L1 IHC、MSI-H/dMMR が臨床導入されているが、より精度の高いバイオマーカー探索が続いている。ICI は、腫瘍免疫に作用する抗体であり、腫瘍免疫サイクルを構成する複数の作用点が影響するため効果予測を困難にしている。現時点で、腫瘍変異量 (TMB)、TCR レパトア、腫瘍浸潤リンパ球 (TIL) などの研究が進んでいる。申請者らは、AMED(課題 ID18980565、20314944)、PRISM 浜本班を通して、肺癌における PD-L1 の発現、腫瘍変異量 (TMB)、腫瘍に浸潤しているリンパ球 (TIL) 等と予後、治療反応性に関して研究してきた。これら AMED、PRISM 事業を基盤とした病理形態学的特徴の解析から、がんの組織像は様々な遺伝子、免疫環境を反映することが確認されており、バイオマーカーの開発対象として有望と考えている。研究代表者らは、科研費基盤研究で、「がん免疫療法効果予測に資する病理学的指標の探索」を行い、がん組織の形態学的特徴量、CD8 陽性 TIL は治療効果、再発、生存期間と関連することを見出し「形態学的バイオマーカー (Morphological biomarker; MBM)」の有用性を提唱している。MBM は人工知能 (AI) によるデジタル画像判定が可能であり、その開発を産学連携して進めてきた。他のバイオマーカーと比して検体品質の影響を受けず、追加侵襲はない等の利点があり、がん患者に優しい医療を実現できる。本研究は、世界初の病理画像 AI 判定機器の実用化を目指している。病理検体は確定診断のために必ず採取されるもので、病理形態像はがんの本質を最もよく反映している。本医療機器の開発を通して、病理デジタル画像バイオマーカー分野を新規に開拓することで、患者への負担が少なく、治療選択に有用な指標を広く提供することできる。また、本研究で開発する AI 画像判定機器は、臓器横断的適応拡大、他の薬剤への応用など、更なる研究開発の発展性がある。

研究の概要

肺癌に対する免疫チェックポイント阻害剤 (ICI) の薬効予測を行う病理デジタル画像 AI 判定機器の開発の一環として、LC-SCRUM レジストリを用いる性能検証のための後ろ向き観察研究である。申請者らは、がんの形態学的特徴は、癌の遺伝子変異状態、免疫微小環境、予後、治療効果と密接に関連することを研究してきた。病理組織デジタル画像を用いて画像解析により抽出した特徴を「形態学的バイオマーカー (morphological biomarker; MBM)」として位置づけ、治療効果予測に応用する技術開発を進めている。これまでに単施設の後ろ向き観察研究を行い、MBM は、既存バイオマーカーである PD-L1 より優れた新規指標となる可能性を報告してきた。MBM は、AI 画像解析技術を応用することで客観的な指標とすることが可能で、追加侵襲の必要はない。さらに、検体の保存や核酸やタンパク質など品質に影響されることがなく、NGS 等による腫瘍変異量 (TMB) に比較して、より多くのがん患者で確実に検索できることは大きな利点である。また、がん組織の形態評価は、がん細胞の異常だけでなく、免疫微小環境を評価することもできるため、多面的解析を行うことができる。このことから、がん免疫に作用する免疫チェックポイント阻害剤の薬効予測には、より包括的指標として高い精度が期待される。本研究では、本邦最大の肺癌レジストリである LC-SCRUM のうち IBIS 試験に登録された症例を対象に、AI 判定機器の予測性能を検証する。また、LC-SCRUM-IBIS は、PD-L1 や遺伝子解析の情報があり、ベンチマークとなる PD-L1 など他の薬効予測指標との比較が容易にできる。本研究では、以下の 2 つの研究開発を通し、MBM-AI 判定機器を医療機器として実用化することを目標とする。(1) LC-SCRUM レジストリを用いた MBM-AI 判定機器の性能検証試験 肺癌患者で ICI 投与例における MBM の治療効果予測性能の多施設観察研究を実施し、単施設研究で開発した AI 判定アルゴリズムをプロトタイプとして、LC-SCRUM-IBIS レジストリ検体を対象に MBM の AI 判定機器の実効性と性能に関して以下の 3 項目について研究を行った。

(1) LC-SCRUM 登録症例を用いる MBM-AI 判定機器の性能検証試験

LC-SCRUM 登録症例は全 971 例で、この中で ICI 投与例 191 例、治療効果のうち SD を除外した 156 例を対象とした。治療効果有り (CR/PR) と無し (PD) に分割し、MBM-AI 判定、PD-L1 IHC および病理医目視評価による指標 (Pathological prediction index, PPI) を用いた薬効予測性能を算出した。その結果、MBM AI モデルでは PD-L1 判定による予測に対して感度が同等で精度の高いモデルを確立することができた。PD-L1 の精度・感度・特異度は 0.48/0.80/0.38 に対して、MBM AI モデルでは 0.69/0.80/0.62 の成績が得られた。一方、病理医目視評価の結果は、最も成績のよい評価者の成績は 0.82/0.83/0.81 と優れていたが、最も低い成績は 0.68/0.5/0.75 で、病理医間の一致度は 0.3~0.5 (Cohen' s kappa 値) と再現性に課題があった。以上の結果より、MBM-AI 判定機器は安定して精度の高い ICI 薬効予測スコアの算出が可能であることが示された。

(2) MBM の AI 判定機器の最適化

MBM AI モデルのプロトタイプ (v1) に対する改良を複数条件 (v2, v3) で試行し、最終的に PD-L1 判定による予測に対して感度が同等で精度の高いモデルを確立することができた。PD-L1 の精度・感度・特異度は 0.48/0.80/0.38 に対して、MBM AI モデルでは 0.69/0.80/0.62 の成績が得られた。

本研究では **Multi-instance learning (MIL)** という学習アルゴリズムをベースとして独自のタイル選択方法を組み合わせて精度良く学習する方法を確立した。実用化に際しては、AI 判定の確からしさをヒト (病理医) が判断できることが求められる。本研究対象での誤判断した要因として、解析対象画像において腫瘍細胞の量や質が不十分な症例が含まれていたが、病理医目視評価でも要因の特定できない例もあった。後者は患者の全身状態など腫瘍以外の臨床要素の影響が大きい可能性はあり、適応には注意が必要であることが示唆された。前者に対しては、AI による対象領域を明示するプログラムを開発し見逃し防止対策を行うことで回避可能と考え、新規にフローに追加した。

以上、MBM-AI 判定機器として薬効予測の精度だけでなく、プログラム医療機器化としての観点からも最適化は完了した。しかし、AI 判定結果の可視化と妥当性評価法については医療機器実用化に向けての大きな課題であり、今後も継続して研究を推進する必要がある。

(3) MBM の AI 判定機器の実効性検証試験のプロトコル確立

実臨床における AI 活用場面には下記 2 通りの案を考えた。すなわち (1) 病理医の目視評価を補助するデジタル画像解析用の医療機器としての活用、(2) 薬剤選択に直結する MBM-AI 判定を行う医療機器としての活用である。アカデミア、企業、PMDA 等関係各所との相談を進めている。

将来展望

今後、本品 (MBM-AI 判定機器) が実臨床に導入されることで、非小細胞肺癌患者におけるバイオマーカーに基づく治療層別の最適化に寄与することが期待される。本品は ICI 治療薬の効果予測を行うものとして開発してきた。開発開始時 (R3) から現時点 (R6) の間で、治療薬の選択枝に大きな変化があり、PD-L1 検査の結果に関わらず ICI 投与は可能となっているが、一方で依然として投与された患者での治療効果は様々であり、適切な薬剤選択が求められている点には変わりはない。また、医療現場からは高額医療としての社会的背景による不均衡も生じている。本品の活用場面として、より治療効果が期待される患者に対して最適な治療薬を提案するとともに、効果が低いと予測される患者については別のより有効な薬剤・治療法への到達を促すことを想定している。この結果から、患者利益に資するとともに社会保障費を適切に再配分する役割が期待される。

Optimization of cancer care is a crucial goal, and effective treatments must be accurately administered to patients. Based on the proceeded research, it has been confirmed that cancer histology reflects various genes and immune environments. We have found that morphological characteristics of cancer tissue, TILs, are associated with treatment efficacy, recurrence, and survival, suggesting the usefulness of "morphological biomarker (MBM). MBM can be digitally determined by artificial intelligence (AI). This research aims to develop and validate the first MBM-AI determination digital pathology medical system to predict the drug efficacy of immune checkpoint inhibitors (ICIs) for lung cancer.

(1) Performance validation study of the MBM-AI judgment instrument using LC-SCRUM registry specimens

A total of 971 registered in the LC-SCRUM-IBIS study which is the largest multicenter non-small cell lung cancer registry in Japan, of which 156 were recruited who treated with ICI and excluded SD. The performance of MBM-AI, PD-L1 IHC, and pathologist-pathological prediction index (PPI) were calculated. As a result, the MBM-AI model showed similar sensitivity and accuracy to the PD-L1 model, with an accuracy, sensitivity, and specificity of 0.48/0.80/0.38 for PD-L1 and 0.69/0.80/0.62 for the MBM-AI model. On the other hand, the results of PPI were excellent (best 0.82/0.83/0.81), with inter-pathologist agreement ranging from 0.3 to 0.5 (Cohen's kappa value, fair to moderate). These results indicate that the MBM-AI judgment instrument can provide a stable and accurate ICI drug effect prediction score.

(2) Optimization of the MBM AI Decision Instrument

We established a method for learning with high accuracy by combining a unique tile selection method based on a learning algorithm called Multi-instance learning (MIL). For practical use, a human (pathologist) is required to be able to judge the accuracy of the AI judgment. The factors causing misjudgment in this study included cases with poor tumor tissue in the selected areas. We have developed a program that indicates the target area by AI to prevent missing the target area. How to assess the AI judgment by human is the significant issue and continued research is still needed.

(3) Establishment of protocols for effectiveness verification testing of MBM's AI judgment devices

The following two proposals were considered for AI utilization in clinical practice. (1) as a medical device for digital image analysis to assist pathologists in visual evaluation, and (2) as a medical device for MBM-AI judgment directly related to drug selection. We are consulting with academia, companies, PMDA, and other related parties.

Future Prospects

The MBM-AI device is expected to contribute to optimizing biomarker-based treatment stratification for non-small cell lung cancer patients when it is introduced into actual clinical practice. The product is expected to suggest the most appropriate therapeutic agent to patients who are expected to have a better therapeutic effect and to encourage patients who are expected to have a lower effect to reach for another more effective agent or therapy. The results of this study are expected to contribute to patient benefit and play a role in appropriately redistributing social security funds.