

総説

外科と AI

獨協医科大学 下部消化管治療センター

石塚 満, 渋谷 紀介, 蜂谷 裕之, 西 雄介, 高柳 雅,
藤田 純輝, 井原 啓佑, 中村 隆俊, 入澤 篤志Artificial Intelligence (AI) による
外科手術の展望

今日、本邦においてもロボット支援下手術はすでに現実のものとなり、臨床現場において、様々な疾患に対してロボット支援下手術が行われるようになってきた。

現時点では外科領域における Artificial Intelligence (AI) の臨床応用はまだまだ一般的ではないが、近い将来にはロボット支援下手術と同様に AI の外科領域への臨床応用も現実となることが予想される。そこで外科領域における AI が果たす役割についてこれから述べていきたいと思う。

AI ナビゲーション手術

消化器外科領域では、現在多くの手術が内視鏡補助下で行われていることから、内視鏡手術で展開される内視鏡画像に AI の画像診断能を付加させることは非常に有用であると考えられる。下部内視鏡検査でモニター上に現れる隆起病変を AI が瞬時に判断し、切除すべき病変かそうでない病変かを判断するように、外科手術においても AI が切るべきものと切るべきでないものを瞬時に判断し、適切な切離予定線を、モニター上に描くことで出血や他臓器損傷のリスクが少ない安全性の高い手術を行うことが可能となる。

国立がん研究センター東病院食道外科からの報告によると、食道癌患者における AI を使用したナビゲーション手術では、AI の術中反回神経の早期認識率は食道外科専門医と遜色ない水準であり ($p=0.26$)、食道外科を専門としない消化器外科医と比べると AI はその認識能において消化器外科医を有意に上回っていた ($p<0.05$)¹⁾。これらのことから AI が反回神経を同定する能力は、すでに食道癌のエキスパートに匹敵するレベルにあるといえる。

彼らに AI の技術提供を行っているのが、国立がん研

究センター東病院から生まれたベンチャー企業 Jmees である。Jmees は医用画像処理・コンピュータ支援外科分野のトップカンファレンスである MICCAI (<https://conferences.miccai.org/2022/en/MICCAI2022-CHALLENGES.html>) のコンペの 1 つである手術映像中の臓器などの位置をトラッキングする精度を競う SurgT (<https://surgt.grand-challenge.org>) に参加し、手術映像特有の情報に基づいた画像処理の手法を用いることで世界 1 位の認識精度を達成している。

外科手術映像は出血や臓器の変形、ミスト、手術器具による遮蔽など一般画像物体とは異なる厳しい条件下であり、堅牢なトラッキングも必要とされることから、今回の結果により Jmees の、AI による手術動画の自動認識技術が世界でも非常に高いレベルにあることが立証された。

従来手術では熟練医が持つ優れた解剖構造の認識力、手術工程の計画力、術中の判断力など全てがブラックボックスであるため、それらを習得するには多くの経験を手探りの中で積む必要があった。実際に外科の専門医さらには内視鏡外科の技術認定医を取得するためには年余にわたり多くの経験とトレーニングが必要不可欠であった。さらに手術中の臓器損傷は、患者の QOL 低下や医療コストの増大に直結することから、Jmees が開発した手術映像をリアルタイムに解析し術者の認識力を支援する AI 手術支援システムは、より安全な手術の施行を実現させた。

さらに進んで、事前学習によりエキスパートの手術を学習し、その手技を習得した AI により、エキスパートの技術と手法を実際の手術において再現することも可能となるであろう。こういった AI によるナビゲーション手術の進歩により、初学者であっても、さらにはいかに不器用な術者であっても AI によるサポートを受けることによりエキスパートと同様の安全で精緻な手術を行うことが可能となるであろう。

先に述べたようにロボット支援下外科手術はすでに現実のものとなり、今日では、都市部から離れた地域における外科医不足や指導者不足を補うための遠隔手術の導入も実現に向けて動き出している。今後ロボット支援下手術に AI ナビゲーションシステムが加わり進化することで、外科医に代わり、ロボットが AI の指示のもと、緻密な外科手術を行う完全自立式手術施行ロボットシステムも近い将来に実現可能になると予想される。かような状況下では現在の外科医の不足、外科医の高齢化、指導医の不足、エキスパートの局在化、長時間労働といった日本の外科医が直面している様々な問題についても一気に解決をみることになるのではないだろうか。

Artificial Intelligence (AI) による予後予測

加えて人工知能による予後予測も現実的なものとなってきた。外科手術において予後を予測することは過剰な外科侵襲を避ける目的でも重要である。予後不良な患者に対する外科侵襲はときとして予後を改善することなく、患者の余命を縮めてしまうこともある。このため術前に得られたデータを用いて患者の予後予測を行うことはこれまでに多々行われてきた。しかしながら多くの臨床背景因子を用いて単変量解析を行い、その結果得られた予後関連因子を用いて多変量解析を行うことで、より予後と強い相関のある臨床背景因子を導き出すことはそれなりの労力と時間を必要とする作業である²⁾。

このためこれらの作業を簡略することができるシステムが渴望されていた。

Prediction One による予後予測

こういった要望に応えるべく Sony から Prediction One (Sony Network Communications) という AI による各種予測を行うソフトウェアが開発され、無料で始めることができる体験版もネット上に公開された。このソフトウェアでは、エクセルシートなどに記録された患者データを入力するだけで、最も予後との関連のある因子を抽出することが可能となった。

最近の医科歯科大学消化器外科からの報告によれば、下部直腸癌における側方リンパ節転移予測に Prediction One を用いたところ prediction model における The area under the receiver operating characteristic curve (AUC) は 0.903 となり、通常の MRI 画像を用いて描出される最大側方リンパ節の短径から導き出されるこれまでの予測能の AUC=0.754 を有意に上回る良好な結果が得られた ($p=0.022$)³⁾。

最新の大腸癌治療ガイドラインによれば⁴⁾、術前または術中診断にて側方リンパ節転移陽性の場合、下部進

行直腸癌患者に対して側方リンパ節郭清を行うことは強く推奨される (推奨度 1・エビデンスレベル C) とされているが、現時点では側方リンパ節郭清を省略できる症例の基準は明らかでないことに加え、側方リンパ節転移の診断基準も確立されていないのが現状である。同ガイドラインによれば MRI による側方リンパ節転移の診断能は短径 10mm をカットオフとするよりも短径 5mm をカットオフとした方が良好であることが示されているが、先に示した AUC の値からも MRI の術前診断能の限界も示されている。

Prediction One を用いた術前側方リンパ節転移診断で正確に側方リンパ節転移の予測が可能になれば、手術時間の延長、術中出血量の増加、術後排尿・性機能障害の増加といった側方リンパ節郭清に伴う合併症を減らすことが可能になり、手術を受ける下部進行直腸癌患者さんにとっても大いなる福音となることであろう。

さらに、米ルイジアナ州立大学やイェール大学、メイヨークリニックなどの研究チームは、大腸癌肝転移 colorectal liver metastasis (CLM) 術後の再発・死亡を予測する機械学習モデルを構築し報告した。

彼らは 2000 年から 2018 年に CLM に対する外科手術を受けた 1,004 件の臨床データを用いて、128 の臨床変数から再発と死亡を予測する機械学習モデルの構築を行った。データ全体として、切除後の生存期間中央値は 47.2 ヶ月、無病生存期間は 19.0 ヶ月で、コホートでの追跡期間中央値は 32.0 ヶ月であった。いずれのモデルも良好な予測力を示し、既存指標を上回る臨床的有効性が示唆されていた。

著者らは「CLM 術後の生存や再発の予測モデルを作成するためのツールとして、機械学習は非常に強力だ」としており、さらなる改良により実臨床で利用可能な予測システム構築が可能であることを強調している⁵⁾。

ChatGPT

さらに最近の話題としては米オープン AI が開発した様々な質問に臨機応変に答える対話型 AI “ChatGPT” (チャット GPT) があげられる。2022 年 11 月の公開から 1 週間足らずで 100 万人以上が利用し、無料でも使えることから、ユーザー数が爆発的に伸びている。日本語に対応し、プログラムのソースコードの添削もできる。

ChatGPT は世界で話題をさらっており、米国時間 1 月 23 日には、米マイクロソフトがオープン AI に数十億ドル (数千億円) の追加投資を発表。米ブルームバーグは投資額が最大で 100 億ドルになると報じた。マイクロソフトは、オープン AI が開発したモデルを自社製

品に展開し、ChatGPTもマイクロソフトのクラウドサービス「Azure (アジュール)」で使えるようにした。ChatGPTでは、大量のテキストデータをAIに与えて学習させる「大規模言語モデル」を採用し、インターネット上の膨大なデータから学習を行い、人間の質問に適切に回答できるようにトレーニングされている。オープンAIによれば、ChatGPTはさらに「強化学習」と呼ばれる段階を経ていて、回答の内容が好ましいものかを人間が評価し、フィードバックすることでモデルを微調整している。

AIによる手術適応

ここまで筆を進めてきて私も実際にChatGPTに臨床上の疑問点をぶつけてみることにした。なぜならばマサチューセッツ総合病院・ハーバードメディカルスクールの研究チームは、米国の医師資格試験であるUSMLEの3つのコンポーネントをChatGPTがパスしたとして、プレプリント論文をmedRxivから公開し、話題となったからである。

以下はその結果である。

「下部進行直腸癌に対して側方リンパ節郭清を行うことで予後の改善が認められるか？」という私の問いかけに対してChatGPTは以下の返答を行った。

「はい、側方リンパ節郭清は下部進行直腸癌の治療として有用である場合があります。側方リンパ節郭清は、リンパ節内の癌細胞を取り除くことで、癌が再発するリスクを減らすことができます。また、側方リンパ節郭清によって、癌が進行していないかどうかを確認することもできます。」

「しかしながら、側方リンパ節郭清が必ずしも予後を改善するとは限りません。癌の進行度や患者の健康状態など、多くの要因が予後に影響するためです。そのため、側方リンパ節郭清が適切な治療となるかどうかは、医師と患者との共同判断に基づいて決定する必要があります。」

なんともわかったかのようなそれでいて曖昧さの残る回答である。

一方でPerplexity AIに同じような質問をしてみるとDoes lateral lymph node dissection improve oncologic outcome of patients with lower rectal cancer? Evidence suggests that lateral lymph node dissection (LLND) in addition to neoadjuvant (chemo) radiotherapy may improve oncological outcomes^{1,2)}. LLND has been shown to reduce local recurrence of patients with rectal cancer and improve survival^{3,4)}. However, the efficacy of LLND without neoadjuvant (chemo)

radiotherapy is still unclear⁵⁾.

と比較的まともな回答が得られ、参考文献までついてくることから、こちらが当惑(Perplexity)することもなかった。

ChatGPTは非常に便利で、どのような展開で話していこうかと悩むような内容のサンプルを即時に作ってくれる。一方で、その内容については誤っているものも多く、情報の取捨選択については利用者が正しく判断できることが必要だと考えられる。先ほど示した例でもかなりあやふやな答えをだしてくるからChatGPTの進言を鵜呑みにすることは避けるべきだと思われる。

よくある説明としては、ChatGPTはコピーキーに例えられる。つまりコピーした鍵を使って鍵を複製していくと次第にずれが生じ、最後には開錠できなくなってしまうことから、利用者は、あくまで情報の“マスターキー”である一次情報にアクセスする意識が必要であるというものである。さらに、AIが提案する文章は誰かのテキストデータを学習させたものであることからChatGPTを仮に論文の共著者とするのであれば、利用されたテキストデータの作成者も共著者に並べられるべきだろう。実際に最近ではChatGPTを共著者とする論文も散見される⁶⁾。

事実、ChatGPTが書いた論文は、学習した情報と具体的な指示内容に限定されるため、オリジナリティに欠けるだけでなく、ChatGPTが用いた情報の正確性を検証する必要もあるとされている。結論として、ChatGPTはテキスト作成の支援に限定して利用されるべきであり、論文の主旨とは関係ないさわりの文章については、ChatGPTに作成を頼むぐらいはリーズナブルな判断だと思われる。しかしながら、提示された文章を反芻・評価して自分の表現に修正することは最低限必要であろう。

AIによる術後管理

術後管理分野においてもAIの予測能に期待が高まっている。ICUせん妄は、集中治療室(Intensive Care Unit; ICU)入室中の成人に起こる急性精神障害で、失見当識や知覚障害、思考錯乱、情緒障害など、多彩な症状を呈する。過活動型のせん妄は症状が強いため医療者にも認識されやすいが、低活動型のICUせん妄はその多くが見過ごされている一方で、患者予後を悪化させる点では過活動型と変わらない点がこれまで問題視されてきた。

ICUせん妄をAIを用いて改善する目的で、ジョンズホプキンス大学の研究チームは、電子カルテで日常的に収集される臨床データおよび生理学的データに機械学習



図 1

夢のスーパーカー“Knight Industries Two Thousand”

Wikimedia Commons より

KITT (Knight Industries Two Thousand) interior

https://upload.wikimedia.org/wikipedia/commons/2/20/Knight2000_ex107.jpg

を適用し、ICUせん妄を予測できるとする研究成果を公表した。

彼らは全米 208 病院、20 万件を超える ICU 滞在データを利用し、この機械学習モデルのトレーニングを行った。彼らの報告によると「静的モデル」と名付けられたモデルは、ICU 入室後 24 時間のデータを用いて、その後 ICU 滞在時にせん妄を発症するかを予測するもので、もう一方の「動的モデル」と名付けられたモデルは経時データから、12 時間後までのせん妄発症を予測するものである。性能検証において「静的モデル」が AUC 0.73、「動的モデル」が AUC 0.85 と良好な値を示し、ツールの臨床的有効性が示唆される結果となった。

著者らは「臨床データおよび生理学的データで学習した機械学習モデルは、ICU せん妄を予測し、時間的な制約のある中でも動的予測をサポートすることができる」と結論づけている⁷⁾。今後は一般病棟での術後管理にも応用が期待される。

最後に

この原稿を書いている間にもテスラが全米で、完全自動運転ソフト (full self driving soft ; FSD soft) の不具合から 36.2 万台をリコールするというニュースが入ってきた。調べてみたところ自動運転にはレベル 1-5 までの 5 段階の定義があり、アダプティブクルーズコントロールや駐車支援などレベル 1-2 の機能を備えた車輛は、現在すでに路上を走っており、レベル 5 は、運転は

もちろん複雑な判断や緊急措置をすべて自動車が実行するレベルで、このレベルになると車輛は人間からの指示は不要となり、ハンドルやフットペダルもなくなってしまふとされている。

つまり AI ソフトウェアによるレベル 5 の車輛はさながらナイトライダーに登場する「ナイト 2000」(図 1) のような車になるわけで、まさに昭和の時代に描かれた夢物語が令和の時代を迎えて俄然現実味を帯びてきたわけである。現在の AI 支援下外科手術支援システムがレベル 1-2 の機能とするならば、AI の進歩により将来的には、レベル 5 にあたる先に述べた完全自立式手術施行ロボットシステムが登場する日も近いのではないかとと思われる。ロボットを使って手術をするのではなく、ロボットが手術を行うのだ。

West-Shibu 888 : 「ドクター、ワタクシガシュジュツヲシテオキマスノデ、ドウゾオヤスマイクダサイ。シュジュツガオワリマシタラオヨビイタシマスノデ。」
私 : 「それじゃ、よろしくね ! ?」

文 献

- 1) Sato K, Fujita T, Matsuzaki H, et al : Real-time detection of the recurrent laryngeal nerve in thoracoscopic esophagectomy using artificial intelligence. *Surg Endosc* **36** : 5531-5539, 2022.
- 2) Ishizuka M, Nagata H, Takagi K, et al : Inflammation

- tion-based prognostic score is a novel predictor of postoperative outcome in patients with colorectal cancer. *Ann Surg* **246** : 1047-1051, 2007.
- 3) Kasai S, Shiomi A, Kagawa H, et al : The Effectiveness of Machine Learning in Predicting Lateral Lymph Node Metastasis From Lower Rectal Cancer : A Single Center Development and Validation Study. *Ann Gastroenterol Surg* **6** : 92-100, 2021.
 - 4) Tomita N, Ishida H, Tanakaya K, et al : Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2020 for the Clinical Practice of Hereditary Colorectal Cancer. *Int J Clin Oncol* **26** : 1353-1419, 2021.
 - 5) Moaven O, Tavorara TE, Valenzuela CD, et al : Machine Learning Models for Predicting the Outcomes of Surgical Treatment of Colorectal Liver Metastases. *Am Coll Surg*. 2023 Jan 19. Online ahead of print.
 - 6) ChatGPT Generative Pre-trained Transformer ; Zhavoronkov A : Rapamycin in the context of Pascal's Wager : generative pre-trained transformer perspective. *Oncoscience* **9** : 82-84, 2022.
 - 7) Gong KD, Lu R, Bergamaschi TS, et al : Predicting Intensive Care Delirium with Machine Learning : Model Development and External Validation. *Anesthesiology* **138** : 299-311, 2023.