

麻酔とテクノロジー

徹底分析
シリーズテクノロジーの
麻酔への応用淘汰と進化の先にある
麻酔の未来

内藤 祐介 15a 新ゴM

近年、麻酔科領域にも人工知能 (AI) やアルゴリズムを中心とした新しい技術が次々と導入されつつある。これらのテクノロジーは、従来のモニタリングや臨床医の判断を補完するだけでなく、臨床の質そのものを変え得る可能性を秘めている。一方で、その有効性や限界を理解しないままでは、せっかくの技術も活かされずに淘汰されることも少なくない。

本稿では、周術期低血圧予測アルゴリズムと区域麻酔におけるAI超音波画像認識を取り上げ、麻酔科におけるテクノロジー活用の現在地と課題を考察する。

13a トラギ/明報 (W2) 17w 12H

16a トラギ/B 17H 新しく登場するモノ・淘汰されるモノ

2020 年以降、生成 AI の登場に象徴される AI ブームは目覚ましいが、この潮流は最近始まったわけではない。振り返ってみると、過去にもその時代の最先端技術を搭載した新しい機器が次々と麻酔科臨床にも導入されてきた。例えば、心電図電極を貼付するだけで心拍出量を測定できる機器や、酸素飽和度センサーのみでヘモグロビン値を推定できる機器など。これらの機器は一部の施設では現在も使用されているかもしれないが、麻酔科臨床に広く普及するには至らなかった。普及しなかった理由は大きく二つに分けられる。

第一は機器側の問題である。精度の不足や実臨床との乖離、他製品で代替可能であること、コストに見合わない、などが普及を妨げる要因となった。第二はエンドユーザーである麻酔科医側の問題である。機器の原理や特性を十分に理解しないまま、その真の有用性や限界点を認識できず、本来の威力を発揮ことなく淘汰されていったのかもしれない。もちろん、麻酔科医は医療機器の専門家ではないため動作原理の深い理解までは必要ないかもしれな

いが、麻酔科専門医試験に麻酔器の始業点検が含まれていることからわかるように、使用する機器の仕組みについて最低限の知識をもつことは求められる。

以下では、現時点で麻酔科領域における「最先端」の技術として注目され、実臨床への普及が進みつつある低血圧予測アルゴリズム Hypotension Prediction Index (HPI) と、区域麻酔で神経などの解剖学的構造物を同定する超音波装置を取り上げ、麻酔科領域のテクノロジー活用について考察したい。

HPI の
考察の前に

HPI は機械学習を用いたアルゴリズムにより低血圧予測を行う機器である。この題材を扱う際には必ず二つの議論が持ち上がる。第一は、卵が先か鶏が先か問題である。すなわち、低血圧に曝露された結果として術後アウトカムが悪化したのか、それとも、もともと術後アウトカムが悪化する素因を術前からもつ患者の脆弱性が術中に低血圧という形で表出したに過ぎないのか、という点である。過去の低血圧研究の多くは後方視的検討であり、研究デザイン上、因果関係を証明できないこと

が背景にある。

第二は、低血圧そのものが悪いのか、心拍出量 (CO) の減少が悪いのかという議論である。周知のように、平均動脈圧 (MAP) は中心静脈圧 (CVP)、体血管抵抗 (SVR) を用いて「 $MAP = CO \times SVR + CVP$ 」と表されるため、MAP 低下はしばしば SVR の低下の反映に過ぎない。逆に SVR が低く MAP が下がっても容認できる場合もある。

誌面の都合上ここでは深く触れないが、この2点はいずれも重要な視点であり、低血圧のみを盲目的に重視する立場には筆者も賛同していない。以下では純粋に HPI と低血圧について述べる。

HPI

HPI は、術中低血圧を「 $MAP < 65$ mmHg が 1 分以上持続する状態」と定義し、その発生を事前に予測することを目的に開発された。HPI は 0 ~ 100 の無次元指数で表され、デフォルト設定のカットオフ値 85 では中央値で約 4 分前にアラートを発する。わずか 4 分前という点に有用性を疑問視する声もある。しかし、ごく短時間の低血圧であっても不良アウトカムと関連することが報告されており¹⁾、この時間的ギャップを HPI で埋める意義は大きい。クリアサイト®を用いた連続血圧モニタリングを基盤とする戦略を、マンシエット管理のみの対照群と比較した研究では、連続測定群で低血圧時間は有意に減少した。ただし、連続測定群でも低血圧曝露を完全にはゼロにすることはできなかった²⁾。これは、アラートが発せられてから、薬

物の準備、投与、作用発現に至るまでの時間が影響していると考えられる。

HPI の原理と精度

HPI は動脈圧波形 (A-line) から得られる特徴を用いる。まず波形から多様な特徴量を抽出し、それらを組み合わせさせて約 260 万通りの特徴を生成する。その中から最も術中低血圧 intraoperative hypotension (IOH) に関連する 23 個の特徴が抽出され、モデル構築に用いられた³⁾。最終的に採用された特徴は非公開だが、ロジスティック回帰モデルであることが公表されている。外部検証では感度 92%、特異度 92%、受信者動作特性曲線下面積 area under the receiver operating characteristic curve (AUROC) 0.97 で、低血圧発生の 5 分前に警報可能と報告された。

モデル作成過程を整理すると、①データ収集・前処理、②特徴抽出、③特徴選択、④モデル学習と検証の 4 段階に分けられる。データは米国の単施設データベース Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC- II) および Edwards Lifesciences 社が収集した ICU・手術室患者記録を合わせた 1684 例が用いられ、うち 1334 例が学習用、350 例が内部検証用に割り当てられた。さらに外部検証用としてカリフォルニア大学アーバイン校で前向きに収集された 204 例が使用された。これらはいずれも 100 ~ 500 Hz の高解像度動脈圧波形を含んでいる。前処理では心拍検出とアーチファクト除去を行い、低血圧 ($MAP < 65$ mmHg が 1 分以上持続) と非低血圧 ($MAP > 75$ mmHg が 20 分以上持続) の時点をラベル付

けた。急激な大量出血など外因性の血圧低下は除外し、漸進的低血圧のみを対象とした。特徴抽出では波形を 1 拍ごとに分割し、収縮期・拡張期など 5 相に区切って時間、振幅、勾配、面積、複雑性、パロレフレックス応答、変動性、スペクトル解析、といった 3022 特徴を抽出している。これをもとに約 260 万の組み合わせ特徴量を生成し、循環力学的要素を幅広く反映させた。特徴選択では、まず ROC 解析で識別能を評価し、AUROC > 0.85 の 51 種類を基本特徴として抽出、そこから性能の高い組み合わせを選定した。

モデル構築にはロジスティック回帰が採用された。これは 2 値分類に適し、出力を確率として解釈できる利点がある。学習用データで最適化されたモデルは 0 ~ 1 の連続値を出力し、これを 100 倍したものが HPI である。学習は患者群を繰り返し分割しながら行われ、クロスバリデーションで性能が評価された。

こうして HPI モデルは、大規模かつ高解像度データの収集、厳密なラベル付け、数百万規模の特徴抽出と選択を経て完成した。ロジスティック回帰という古典的な統計学でも馴染み深い手法を用いることで解釈性の高い形式を保ちつつ、多次元の循環動態情報を圧縮し、臨床に有用な予測指標として成立した点が特徴である。

12a トラギ/B/OI (以下同)

実際の使用感と注意点

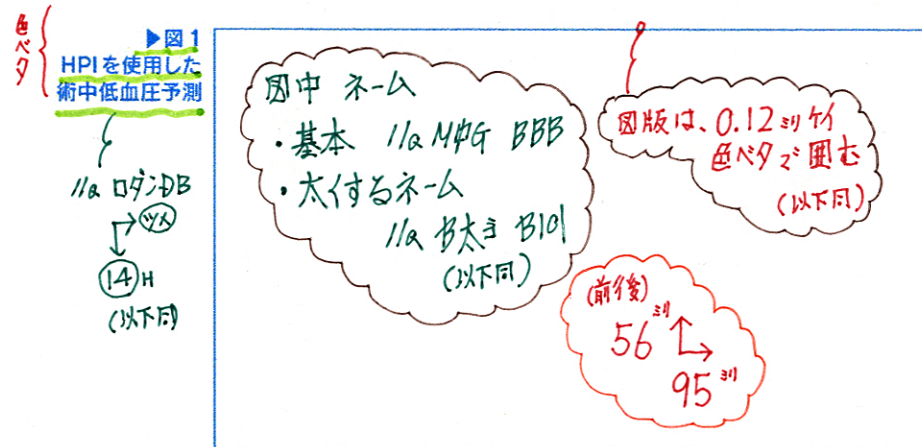
図 1 は HPI を使用した自験例である。MAP が低下傾向を示すと HPI が上昇する様子が明瞭に確認できる。一方で、アラートが鳴動するのは MAP 67 mmHg といった「厳密には低血圧の基準を満

10a トラギ/明報 (W3) 12H
NAITO, Yusuke
奈良県立医科大学 麻酔科学教室

徹底分析
シリーズ

麻酔とテクノロジー

14a ロダDB (YX)



たしていないが、誰がみても低血圧が予測できる段階」であることも多い。実際に図をみれば、MAPの下降とHPIの上昇はまるで鏡のように一致して変動しており、つまりは、MAPのトレンドに注意していればHPIがなくとも低血圧を予測できる、と感じなくもない。

この発想をもった研究者はすでに存在し、この疑問に対する回答も示されている。後方視的研究ではあるが、2施設における29万2025件の手術患者血圧データを解析し、一般的に低血圧と関連するとされる1回心拍出量変動(SVV)や Δ MAPといった指標とHPIの予測精度を比較した。その結果、SVV、脈圧(PP)、心拍数(HR)、 Δ MAPといった単一指標よりもHPIのほうが高精度に低血圧を予測できることが示された⁴⁾。

さらに、HPIを搭載しているAcumen(Edwards Lifesciences社)は、HPIに加えてdp/dtやEaDynといった指標も同時に提供する。dp/dtは心収縮力の一部を、EaDynは後負荷の一部を反映する指標である。これらを組み合わせることで循環動態をより深く把握でき、低血圧への介入が必要かどうか、必要であれば輸液、血管収縮薬、強心薬のいずれを選択するべ

きかといった判断に役立つ補助情報となる可能性がある。

神経ブロックでの
AIの可能性

画像認識はAIの得意分野の一つである。特に胸部X線画像を用いた判別精度は高く、肺炎、無気肺、心拡大、腫瘍、結節、胸水、気胸といった14種類の病変について放射線科医との読影精度を比較したところ、14病変中10病変で同等の結果が得られた。さらに、平均読影時間はAIが1.5分であったのに対し、放射線科医は240分を要していた⁵⁾。このように画像診断領域において、AIはすでに専門家とほぼ同等の性能を、より短時間で発揮できる段階にまで発展している。麻酔科に関連する領域としては、術前の顔写真から挿管困難を予測するモデル⁶⁾が研究レベルで報告されている。

そして、麻酔科領域でAIによる画像診断が最も普及しているのは、区域麻酔領域であると考えられる。図2はGEヘルスケア・ジャパン社の超音波装置が、腕神経叢、大腿神経、坐骨神経をリアルタイムに着色表示している画像である。プローブの動きに合わせてリアルタイムに神経を同定すること

が可能である。このように、末梢神経ブロックでは神経を同定し、構造物をセグメンテーション(領域塗りつぶし)やバウンディングボックスで表示する手法が多い。一方で、脊髄も膜下麻酔や硬膜外麻酔では、椎間レベルや椎間腔の同定、皮膚から硬膜外腔までの距離推定が中心となる。

画像同定の原理 色ベタ

AIはどのように神経を同定しているのだろうか。以前は「特徴抽出+分類」が主なアルゴリズムであったが、2016年を境に、畳み込みニューラルネットワーク(convolutional neural network, CNN)を中心とする深層学習(deep learning)が主流となっている。CNNは画像や時系列データ解析に特化したニューラルネットワーク構造である。ここでは、簡略化しつつその原理に簡単に触れておきたい。

そもそもニューラルネットワークとは、人間の脳の神経回路を模倣して設計された計算モデルであり、機械学習の一分野として広く用いられる。脳の神経細胞が電気信号を受け取り、伝達・変換して複雑な処理を行う仕組みを数学的に抽象化したものである。図3に示すように、基本構成は「入力層」「隠れ層」「出力層」であり、各層は多数のノード(node, 結合点)からなる。ノード間は「重み」と呼ばれるパラメータで結合され、入力データは重み付けされて次層に伝達される。各ノードは活性化関数によって新しい値に変換され、この操作を層ごとに繰り返すことで、単純な入力から複雑なパターンや関係性を抽出できる。

ニューラルネットワークの学習は誤

徹底分析シリーズ ◆ 麻酔とテクノロジー

スミ80%・スミ80%
6/11 5.5
13a 見出し HB 31
70 3/4

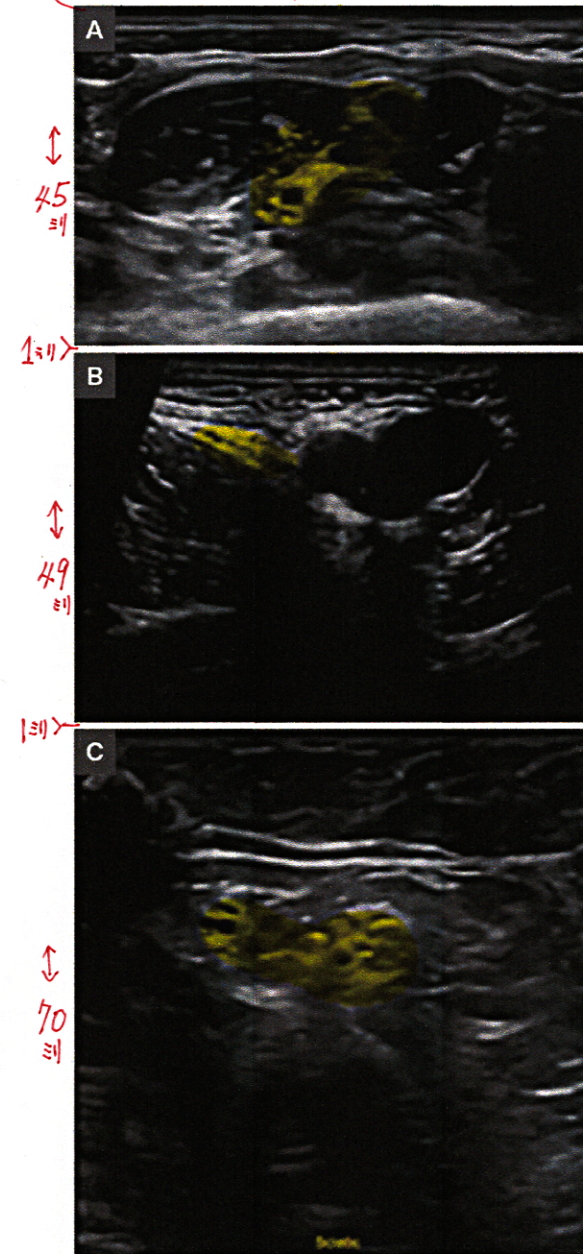
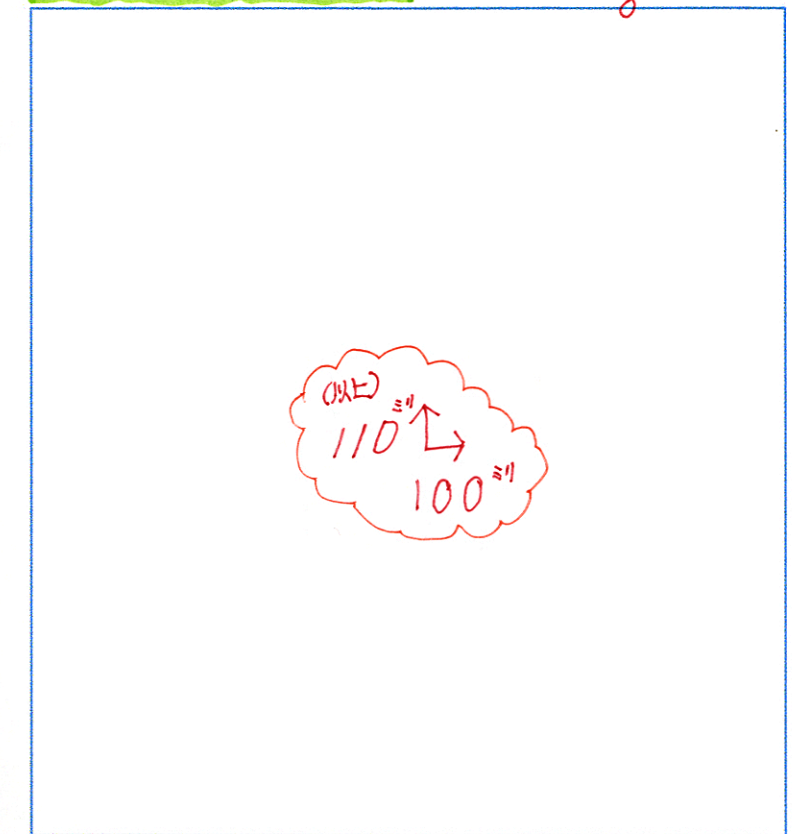


図2 VenueTMシリーズによる神経の同定
(画像提供: GEヘルスケア・ジャパン社) 1/a M4g BBB (YX)
A: 腕神経叢
B: 大腿神経
C: 坐骨神経

図3 ニューラルネットワークの概念図



差逆伝播法(backpropagation)にもとづく。入力値から得られた予測と正解との誤差を計算し、その誤差を減らす方向に重みを調整する。この過程を繰り返すことでネットワークは特徴を自動的に学習し、分類や予測が可能になる(図3)。

あらためて図3を見直すと、従来の全結合ネットワークでは、画像1枚の全ピクセルがまず一次元データに変換され入力層に入る。中間層はすべての入力との結合を計算するため、画像左上のピクセルと右下のピクセルの

ように位置的に関連の薄い結合も含まれ、無駄が多い。この「全層結合」は網羅的ではあるが、計算コストが高いという欠点があった。

CNNはこれを解決するため、画像を二次元配列のまま入力する。小さな領域(例: 2×2ピクセル)単位で畳み込み処理を行うことで、空間構造を保ったまま効率的に特徴を抽出できる。また、従来型ニューラルネットワークでは中間層が多数形成されても各層の役割が明確でなく、チューニングが困難であった。一方、CNNでは浅い層

はエッジ・線・角・模様といった低レベル特徴、中間層は形のパーツや局所的陰影、深層は臓器の輪郭や病変の浸潤影といった抽象的特徴を段階的に抽出する。こうして学習速度が大幅に向上し、高精度な画像認識が可能になった。

問題点 色ベタ

商用利用されている多くの超音波装置では、目的とする神経の同定率が99%程度とされている。一聞すると極めて高精度のように思えるが、それが臨

39
床的に妥当な水準かどうかは慎重に考
える必要がある。一般に画像識別の精
度評価は二つに大別される。第一は
Dice 係数 dice similarity coefficient
(DSC) や IoU といった画素単位の一
致度評価であり、第二は正解率・感度・
特異度などの全体的な分類指標である。

Dice 係数とは、AI によるセグメンテ
ーション結果と「正解 (ground truth)」
の重なり具合を数値化した指標である。
具体的には、二つの集合 A (AI の予測)
と B (正解ラベル) に対して、

$$\text{Dice}(A, B) = 2|A \cap B| / (|A| + |B|)$$

12.5a
で計算される。Dice 係数は 0 ~ 1 の範
囲を取り、1 に近いほど一致度が高い。
一般的に Dice 係数は 0.7 未満では精
度不足、0.7 ~ 0.8 で概ね許容レベル、
0.8 ~ 0.9 で実用可能レベルとされる。

ただし、許容可能な基準は手技の危
険性に依拠して判断されるべきである。
例えば、大腿神経の同定率が Dice 係
数で 0.9 であったとしても、動脈や静
脈と認識した画像が重なっている場合
は許容できない。このように、精度が
高いことと安全性が高いことは必ずし
もイコールではない。

さらに実臨床では、報告されている
数値より低い精度となる可能性が大い
にある。現時点で多くの評価は模擬環
境にもとづいており、臨床的な一致率
を示した研究は限られている。高い
BMI、体内水分量の差、局所麻酔薬を
注入した後の画像変化、解剖学的バリ
エーションなど多様な実臨床での一致
度は低下するだろう。また、臨床アウ
トカムそのものを評価した報告もまだ
存在しない。

麻酔科医の真の目的は、教育効果の

向上、針外傷リスクの軽減、少ない局
所麻酔薬でより効果的な鎮痛を提供す
ることであり、構造物を同定すること
は手段でしかない。今後、これらのテ
クノロジーを使用することにより臨床
アウトカムが実際にどのように変化する
のか、評価をしていく必要がある。

新しく登場するヒト・ 淘汰されるヒト?

研修医と進路について話をしていると、
しばしば遭遇するのが「将来、医者の
仕事が AI に奪われてしまうのではない
か」という不安の弁である。そこで
思い浮かぶのは、筆者が入局した当時、
先代の教授が「自分が研修医だった頃
はいまのような麻酔器はなく、用手換
気で術中管理をしていた」と語ってい
たことである。その時代には、バッグ
の硬さからコンプライアンスを推測し
たり、口唇の色から酸素飽和度を類推
し、呼吸終末陽圧 (PEEP) や換気量
を微調整したりしながら最適な換気を
提供していたのであろう。用手換気が
当時の麻酔科医のアイデンティティに
どれほどの比重を占めていたかは知る
由もないが、少なくともこのルーチン
はやがて機械によって「奪われた」こ
とになる。

しかし、現在、われわれは麻酔器
という文明の利器を手に入れたこと
で、用手換気を凌駕する安全性、正確
性、持続性を獲得し、麻酔科医のタス
クを別の方向へシフトさせることで周
術期医療の質を向上させることができ
た。そして、麻酔器が登場しても、先
人たちの多くは麻酔科医であり続けた。
それは麻酔器の仕組みを理解し、機械
の一步先を見据えた管理を実践してい

たからである。

したがって、研修医に対する筆者の
答えはこうである。

「麻酔科医の役割は将来、大きく変わ
っているかもしれない。でも、テクノ
ロジーを理解し、リテラシーを身につ
けた人は 50 年後も麻酔科医と名乗っ
ていると思う」

本稿が、麻酔領域に普及しつつある
テクノロジーの理解とその活用に少し
でもつながれば幸いである。

文献 13a 見逃す MB 31
色ペン

1. Walsh M, Devereaux PJ, Garg AX, et al. Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery : toward an empirical definition of hypotension. *Anesthesiology* 2013 ; 119 : 507-15.
2. Maheshwari K, Khanna S, Bajracharya GR, et al. A randomized trial of continuous noninvasive blood pressure monitoring during noncardiac surgery. *Anesth Analg* 2018 ; 127 : 424-31.
3. Hatib F, Jian Z, Buddi S, et al. Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. *Anesthesiology* 2018 ; 129 : 663-74.
4. Davies SJ, Vistisen ST, Jian Z, et al. Ability of an arterial waveform analysis-derived hypotension prediction index to predict future hypotensive events in surgical patients. *Anesth Analg* 2020 ; 130 : 352-9.
5. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis : a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med* 2018 ; 15 : e1002686.
6. Wilk M, Pikiewicz W, Florczak K, Jet al. Use of artificial intelligence in difficult airway assessment : the current state of knowledge. *J Clin Med* 2025 ; 14 : 1602.