

別添 3

厚生労働行政推進調査事業費補助金 (地域医療基盤開発推進研究事業)

I. 総括研究報告書

AIを活用したがんゲノム医療の専門家育成指導事例集とカリキュラム案 (H30-医療 - 指定 - 008)

研究代表者 横山 和明

1. 研究要旨

近年の遺伝子解析技術や人工知能(AI)等の情報通信技術(ICT)の顕著なる躍進は、AI を活用し、遺伝子情報に基づいた医療の実践（以下、AI を活用したゲノム医療）を実現可能な領域まで押し上げている。その先進モデル事例として、研究代表者らにより東京大学医科学研究所で推進されている Watson for Genomics(WfG; IBM 社)を用いた遺伝子変異に対応する治療薬の探索が挙げられる。AI を活用したゲノム医療はがん診療をはじめとするさまざまな医療の現場に革新的な変化をもたらすと考えられるが、現状ではその実装には種々の課題がある。特に、この様な先進的医療を担う人材、特に医師の育成に関しては、教育基盤をどのように整備し、どの時期に教育を行うべきかという課題は、十分な議論がなされておらず可及的速やかに取り組むべき喫緊の課題といえる。

そこで、本研究課題「AI を活用したゲノム医療推進に係る人材育成に関する萌芽的研究」では、AI を活用したゲノム医療を推進する専門家を育成するためのカリキュラムの開発と、その教育実施方法を確立するための方策を模索するため、以下の手順で萌芽的研究を行った。

H30 度から、当院での AI を活用した血液がんにおけるゲノム医療の診療研究の過程を、2 年かけて本研究の研究協力者である医学博士課程在学中(卒後数年程度)の医師 3 名が、研究代表者、分担者らの指導のもと、下記の流れ図に従った On-the-Job Training (OJT)を実際に体験し、学習経験を蓄積した。今回作成した「AI を活用したがんゲノム医療の専門家育成指導事例集とカリキュラム案」では、昨年度末の本研究の中間報告に引き続き、研究協力者らが学習経験を 1 年間さらに蓄積した事で明らかとなった問題点や指導の要点も加えて論点整理を集中的に行った。本研究では、OJT を通じて明らかになった、AI 等の ICT を活用したゲノム医療に係る人材育成の実情、ゲノム解析 AI の仕組み、その特徴を活かした活用法、結果の解釈における盲点について、具体例を挙げながら報告する。また、指導の要点から明らかになった、専門家育成のカリキュラム案もその一部を報告する。

研究代表者

横山 和明 東京大学・医科学研究所 附属病院 血液腫瘍内科 助教

研究分担者

井元 清哉 東京大学・医科学研究所 健康医療データサイエンス分野 教授

古川 洋一 東京大学・医科学研究所 臨床ゲノム腫瘍学分野 教授

湯地晃一郎 東京大学・医科学研究所 国際先端医療社会連携研究部門

特任准教授

はじめに

2007 年に開発された次世代シーケンス技術(NGS)により、短時間で大量に DNA 配列を電子情報化することが可能となり、シーケンスにかかる費用が急激に低減化された。かつてのヒトゲノム解析計画では、30 億文字からなるヒトゲノム情報の電子化には 1990-2003 年の 13 年、27 億ドルを要したものの、現在では 1 日以内、1000 ドル以下で全ゲノムシーケンスを得ることが可能である 1)。今や全ゲノムシーケンスデータは 1,000 ドル程度で手に入る様になった。この様な追い風もあり、NGS をがんの臨床に活用する取り組みが世界中で加速している。NGS で解析した患者検体のデータを、スーパーコンピュータを用いて解析すると、パネル検査で数百以上、全エクソンや全ゲノムシーケンスでは千一数十万個見つかる。この情報を、「学術研究」として実施された NGS の結果、即ち学術文献や変異を収載したデータベース、ガイドラインを参考に、診療に活用する事、具体的には変異の結果を基に精密な診断を行い、効果の期待できる薬剤を選び、最適な治療方針を決定する「プレシジョン・メディスン」を実践する事が次の我々のゲノム医科学におけるミッションである。そのためには、膨大な学術論文などの文献情報やデータベースなどを

参考に、個々の症例において、膨大な変異や遺伝子発現データなど NGS 結果から、癌特異的な変異の一つ一つ（図 1. シングルバイオマーカー）を抽出する事や、シグナル伝達経路等の癌に特異的な複数の遺伝子発現の変動（図 1. 遺伝子発現、パスウェイシグネチャ）を抽出する必要がある。この作業はゲノム変異の臨床翻訳や解釈(キュレーション)と呼ばれる。この作業が「プレシジョン・メディスン」実践におけるボトルネックとなっている。この作業に、人工知能(AI)を活用する事による効果は計り知れない。その先進モデル事例として、研究代表者らにより東京大学医科学研究所で推進されている Watson for Genomics(WfG; IBM 社)を用いた遺伝子変異に対応する治療薬の探索が挙げられる。AI を活用したゲノム医療はがん診療をはじめとするさまざまな医療の現場に革新的な変化をもたらすと考えられるが、現状ではその実装には種々の課題がある。特に、この様な先進的医療を担う人材、特に医師の育成に関しては、教育基盤をどのように整備し、どの時期に教育を行うべきかという課題は、十分な議論がなされておらず可及的速やかに取り組むべき喫緊の課題といえる。そこで、本研究課題「AI を活用したゲノム医療推進に係る人材育成に関する萌芽的研究」では、AI を活用したゲノム医療を推進する医師を育成するた

めのカリキュラムの開発と、その教育実施方法を確立するための方策を模索するため、以下の手順で萌芽的研究を行った。

H30 度から、当院での AI を活用した血液がんにおけるゲノム医療の診療研究の過程を、2 年かけて本研究の研究協力者である医学博士課程在学中(卒後数年程度)の医師 3 名が、研究代表者、分担者らの指導のもと、下記の流れ図に従った On-the-Job Training (OJT)を実際に体験し、学習経験を蓄積した。今回作成した「AI を活用したがんゲノム医療の専門家育成指導事例集とカリキュラム案」では、昨年度末の本研究の中間報告に引き続き、研究協力者らが学習経験をさらに蓄積した事で明らかとなった問題点や指導の要点も加えて議論による整理を集中的に行った。本研究では、OJT を通じて明らかになった、AI 等の ICT を活用したゲノム医療に係る人材育成の実情、が知っておくべき、ゲノム解析 AI の仕組み、その特徴を活かした活用法、結果の解釈における盲点について、具体例を挙げながら報告する。最後に専門家育成のカリキュラム案に関しては、OJT を通じた指導の要点から明らかになった、特に重要と思われる項目の一部(概要版)と、実際に学修する際に必要なキーワードも含め記載した(詳細版)を最後に記載する。

1. がんのドライバー変異

遺伝形質を規定し、mRNA に転写される領域を遺伝子と呼ぶ。ヒトゲノム上には約2万個程度の遺伝子があると考えられているが、それはゲノムの内わずか2%程度に過ぎない(PMID: 22304912、PMID: 11237011)。遺伝子変異には、生殖細胞系列変異と体細胞変異がある。生殖細胞系列変異は親から受け継ぐ先天的な変異であり、体を構成するすべての細胞に見られるDNAの変異である。一方、体細胞変異はヒトが生きていく過程において後天的に獲得したDNAの変異である(PMID: 11237011)。がんは遺伝子の病気であり、その蓄積により正常な細胞ががん化すると考えられるが、変異はそのがん化における役割からドライバー(運転手)変異とパッセンジャー(乗客)変異に区別される2)。

ドライバー変異とは、遺伝子異常ががん細胞の増殖や生存にアドバンテージを与える、つまりがん化に直接関与している変異であり、パッセンジャー変異とは、がん化には関係していない変異のことである(PMID: 22304912、PMID: 11237011)。

次世代シーケンサー (NGS) の実用化により、多数の患者検体を用いてドライバー変異を網羅的に解析する「学術研究」が多くのがん種で行われた。そ

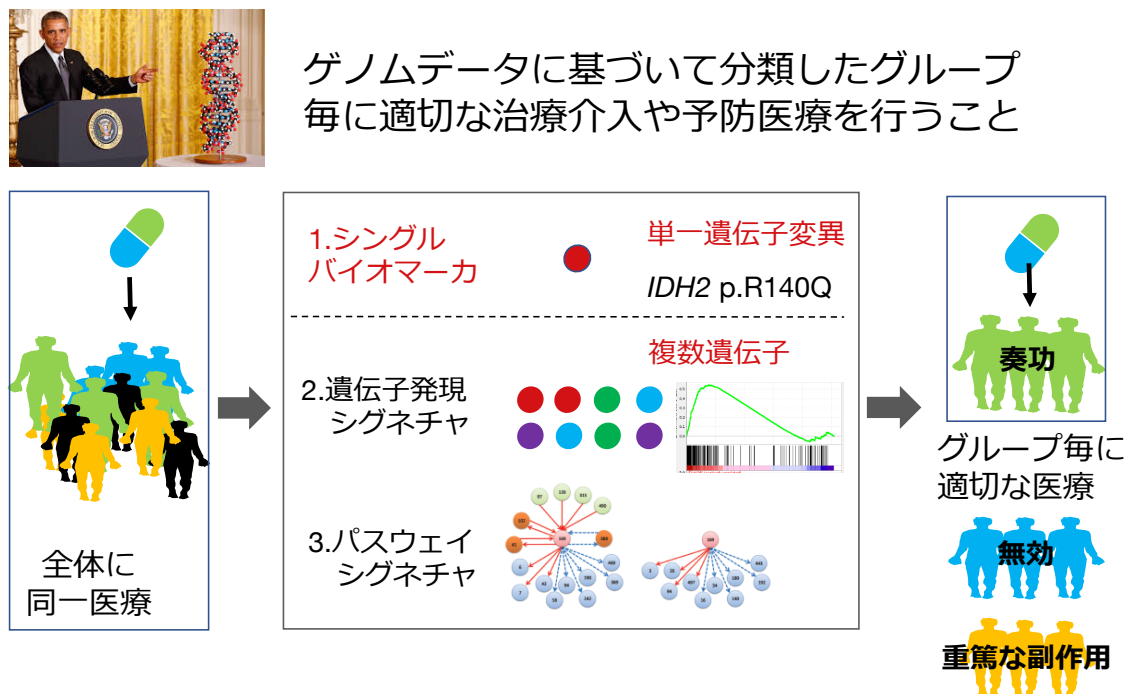
の結果多くのがん種でゲノム情報基盤が確立してきた。それらの膨大な変異のリストはがんにおける体細胞突然変異カタログ(COSMIC) (<http://www.sanger.ac.uk/genetics/CGP/cosmic/>)を始めとして公共のデータとして提供され、一般にも利用可能である。

2. 臨床シーケンスとそのボトルネック

2015 年1 月、当時のオバマ米国大統領による一般教書演説の中で “Precision Medicine Initiative⁴)” というプロジェクト（現在は“All of Us Research Program”と改称）(PMID 25635347)が発表されて以来、この“プレシジョン・メディスン”という言葉は今後あるべき医療のキーワードとして普及した。その言葉の意味を補足すると、「NGS 等の最新技術を用いてゲノム情報、環境要因、ライフスタイルの精密な分析を行い、そのデータを基に患者やその予備軍をグループ分けしてグループ毎に適切な治療介入や予防医療を行うこと」、つまり従来提唱された個人の個別化医療を、グループ毎に推進する事、と考えられる（図1）。実際には、NGSを用いた個人のゲノム情報解読

を核とする新しい医療を米国では“プレジジョン・メディスン”という名称で統一する方向性が提唱されているが、これは“臨床シーケンス”そのものであり、その主たる対象疾患はがんや遺伝性疾患である。NGSにより高速に大量の変異データは出てくるようになったが、これを大量の医学論文、データベース、臨床試験情報などにつきあわせながら臨床的な解釈、翻訳を行うのは非常に労力を要する作業である。例えば現在、代表的ながんの体細胞変異データベースCOSMIC version 903)では、26,829文献から抽出した、1,412,466検体における蛋白をコードする領域の変異として29,519,920変異という天文学的な変異情報が収載されている。さらに、文献情報としてはNIHのPubMed（医学・生物系論文の要旨データベース）上には、がんに関する論文だけでも2018年度には48万件の論文が登録され、2019年1月時点で475万件の論文が登録されている。その数は指数関数的に増え続けている。これらの増え続ける論文やデータベースを参照することはもはや人間の能力の限界を超えている。この一人一人の患者dataの解釈こそが臨床シーケンス推進の上でボトルネックとなっていると言って良い。

図.1



3. ゲノム解析支援特化型AI: Watson for Genomics

人工知能（Artificial Intelligence : AI）とは、人間が行う「知的ふるまい」の一部を、プログラムを用いて人工的に再現したものを指す。特にデータから学習することで、適切な知的ふるまいを人工的に実現するプログラムを機械学習と呼ぶ。人間は、経験から学ぶことによって知的ふるまいが行えるようになる。例えば、猫や犬など何度も動物を見ることによって、動物の種類を見分けられるようになる。機械学習も同じで、写真データから学び、動物を識別でき

るようになる。機械学習をさらに発展させた手法がディープラーニングである。その最大の特徴は、画像認識等、認識に必要な特徴量を人が指示しなくとも AI が自ら特徴量を抽出することが可能となった事である。ディープラーニングは、画像や音声認識、さらには自動運転技術と、さまざまな分野への応用が始まっている。医療分野でも、複雑な医学データのパターン分析や、診断治療の支援ツールとして AI が注目されている。実際に米国では、2018 年 IDx-Dr(糖尿病性網膜症)等の AI 技術を用いた画像診断支援アルゴリズムを FDA が承認した。このように、医療における AI の活用は急速な勢いで進んでいる (PMID: 30617339)。この流れはゲノム解析にもおよんでおり、ゲノム解析支援に特化したクラウドサービスとして代表的な AI は Watson for Genomics(WfG)である(<https://www.ibm.com/jp-ja/marketplace/watson-for-genomics/resources>)。WfG は変異リストのスコア化を行い重要度の高い遺伝子変異、および標的となるパスウェイの推測による最適な薬剤のリストや臨床試験の情報をレポートとして作成してくれる。WfG の解析アルゴリズムは企業秘密で詳細は公開されていないが、研究代表者らは以下の様な、機械学習をベースにしたプログラムであると考えている (図

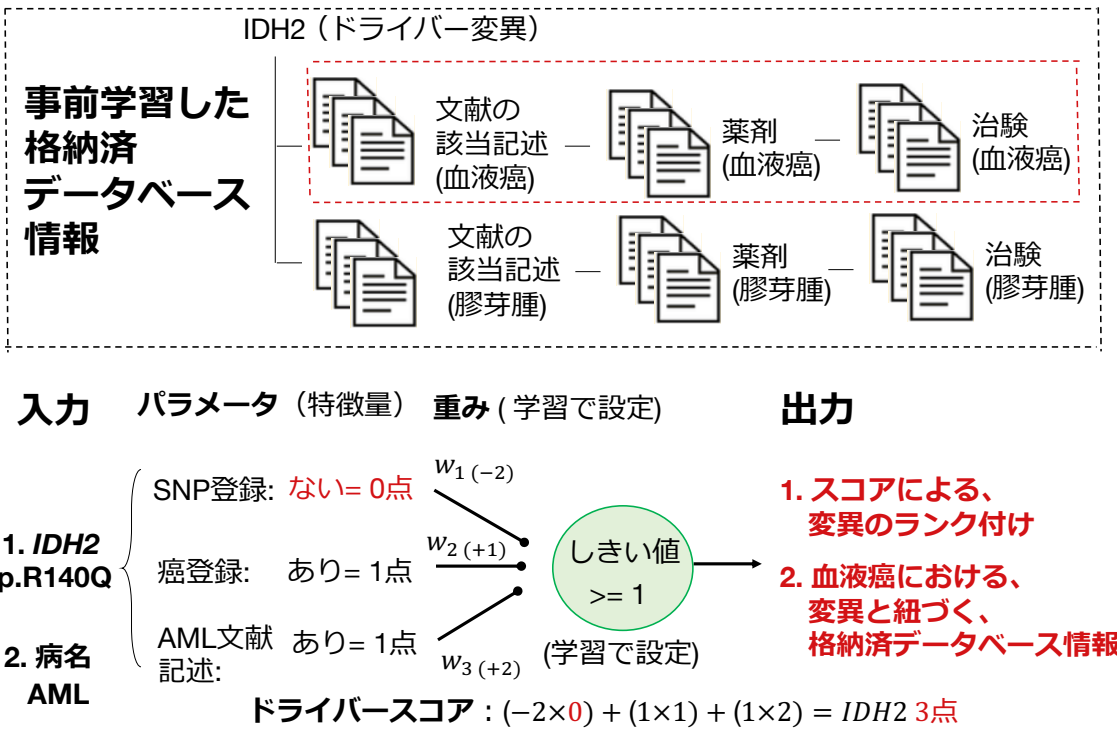
2) 。ドライバー変異としてイソクエン酸脱水素酵素 2(IDH2)遺伝子の p.R140Q 変異を持つ 60 才の AML 男性例を用いて説明する。WfG に入力する内容は、性別、病名、年齢、変異データである。変異データとは VCF(variant call format)と呼ばれる、一行に一変異のデータを記述したデータであり、その具体的な内容は、シーケンスデータをヒトゲノムリファレンス配列にマッピングしたとき、リファレンス配列上の塩基とそこにマッピングされたシーケンスデータ上の塩基などの情報である。入力された IDH2 変異は、WfG が自然言語処理により、キュレーターと学習した格納済みデータベース（コーパスと呼ばれる）を参照しながら、事前学習を基に決定した種々のパラメーターとその重み付けを基に、ドライバー変異である確率（ドライバースコア）が与えられる仕組みである。重みを与えるパラメーターは例えば以下の様なものである。1) IDH2. p.R140Q が SNP データベースに登録があるか、2) 当該変異が癌のデータベース自体に登録の根拠があるか、3) 入力病名の遺伝的背景において、IDH 2 遺伝子および変異が重要なドライバーとしてエビデンス（文献での記述）があるか、4) タンパク質の機能予測プログラムで p.R140Q のアミノ酸変異がもたらすインパクトを予測した場合、変異のイン

パクトがあるか等である。これらのスコアリングシステムを用いて入力変異情報のスコアリングを行い、例えばある閾値以上（簡便の為、図2では3点以上とした）を満たし、結果が上位のもの（例えば上位50個）の変異情報が変異プロファイルにドライバー変異候補として出力される。次に、これらのドライバー遺伝子候補が list up されると、もしそれに付随して学習した関連情報があれば、それらの情報も紐付けて出力する仕組みである。具体的には、1)文献の疾患における該当記述、2)文献における当該変異の記述やデータベースにおける情報、2)変異と紐づいた薬剤情報である。本例の場合、薬剤情報としては、直接のパスウェイ（IDH2）を標的とした阻害剤である FDA 承認薬の enasidenib や、それに関連した、治験データベース 7)に登録のある米国の治験の情報である。出力時には、入力した内容の 60 才男性 AML という内容でさらに出力内容が filtering される。具体的には治験の eligibility において、60 才 AML 男性、対象疾患が AML で entry 可能な治験の情報のみが表示される。次に、もしシグナル伝達における下流の経路など間接的なパスウェイも学習している場合、それを標的とした薬剤とその紐付けられた情報も出力される。例えば、IDH2 変異を持つ解析例(図3)では、AI が学習後には、学習前に見られ

なかった output として IDH 2 変異と紐付けられた間接的なパスウェイ target として bcl2、そしてそれを標的とした bcl2 阻害剤 Venetoclax の治験情報が出力されている。これは IDH2 変異を持つ AML において、IDH1 および 2 の変異が AML で BCL-2 依存性を生じさせ、bcl2 阻害剤の有効性を示唆する既報を学習したからであり (PMID: 25599133、PMID: 27520294)、この格納済みデータベース情報を基に AI が結果を出力したものであると考えられる。

【図 2】

図 .2



【図 3】

図 .3

学習前 AI output:

1	Actionable Alterations	0	FDA Approved for Myeloproliferative Neoplasm	0	Therapies with Clinical Trial(s)	1	FDA Approved for Other Indication(s)
IDH2	R140Q	-	-	-	-	3A	Enasidenib

学習後

3	Actionable Alterations	0	Approved for Chronic Neutrophilic Leukemia	3	Therapies with Clinical Trial(s)	1	Approved for Other Indication(s)
IDH2	R140Q	-	-	Phase 1 Trial 3B Venetoclax	-	2B	Enasidenib

2. 研究目的・方法

医科学研究所における AI を用いた診療支援において、AI はがん細胞のゲノムシーケンス解析を行い、同定されたゲノム変異の臨床的解釈、すなわち「結果の解釈」のプロセス(図 4 赤字部分)で活用されている。通常、シーケンス解析結果の解釈には、一人一人の患者のがん細胞に生じた膨大な数のゲノム変異を、膨大な医学論文、データベース、臨床試験情報等と照らし合わせながら臨床的に解釈する必要がある。これは多大な労力を要する作業である。このプロセスにおいて、理論上は知識量・文献検索量に制約のない AI を活用

する事により、これまでに蓄積されている知識を効率的に医師に与えることができる。これにより診療の質の向上が期待できる。しかし、現状ではこのような先進的医療を担う人材育成に対応した教育基盤をどのように整備し、どの時期に教育を行うべきかという課題は、十分な議論がなされておらず可及的速やかに取り組むべき喫緊の課題と言える。

そこで、本研究では、AI を活用したがんゲノム医療の先進モデル事例として WfG を取り上げ、AI を活用したゲノム医療を推進する医師を育成するためのカリキュラムの開発と、その教育実施方法を確立するための方策を模索するための萌芽的研究を行った。具体的には、まず本研究の研究協力者である医学博士課程在学中(卒後数年程度)の医師 3 名が、H30 年度から H31 年度、そしてそれに引き続き H31 年度から令和 2 年度 3 月まで、計 2 年間を通じて下記流れ図に従い OJT を経験し、各症例について綿密なディスカッションを重ね、臨床的に有意な結果を主治医にフィードバックするという経験を蓄積した。そこで昨年度の中間報告に引き続き、学習経験を通じて明らかとなった問題点や指導の要点を研究代表者と分担者らが集中的にディスカッションして整理した。

この指導経験の中で明らかになった、AI を活用したがんゲノム医療に携わる専門家育成における指導ポイント具体例を挙げつつポイントを絞って提示する。

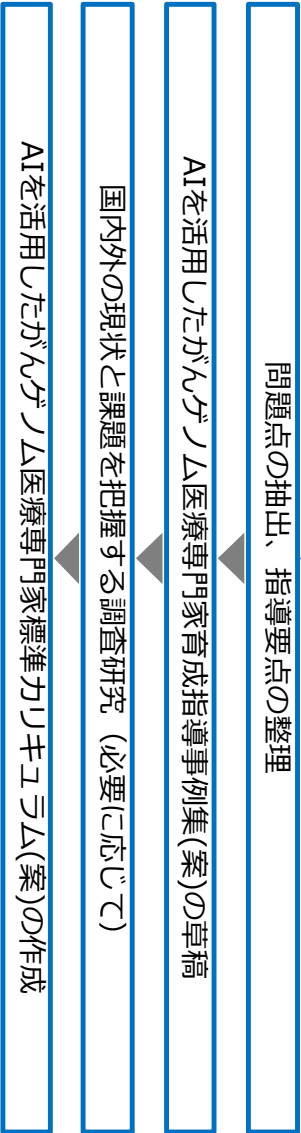
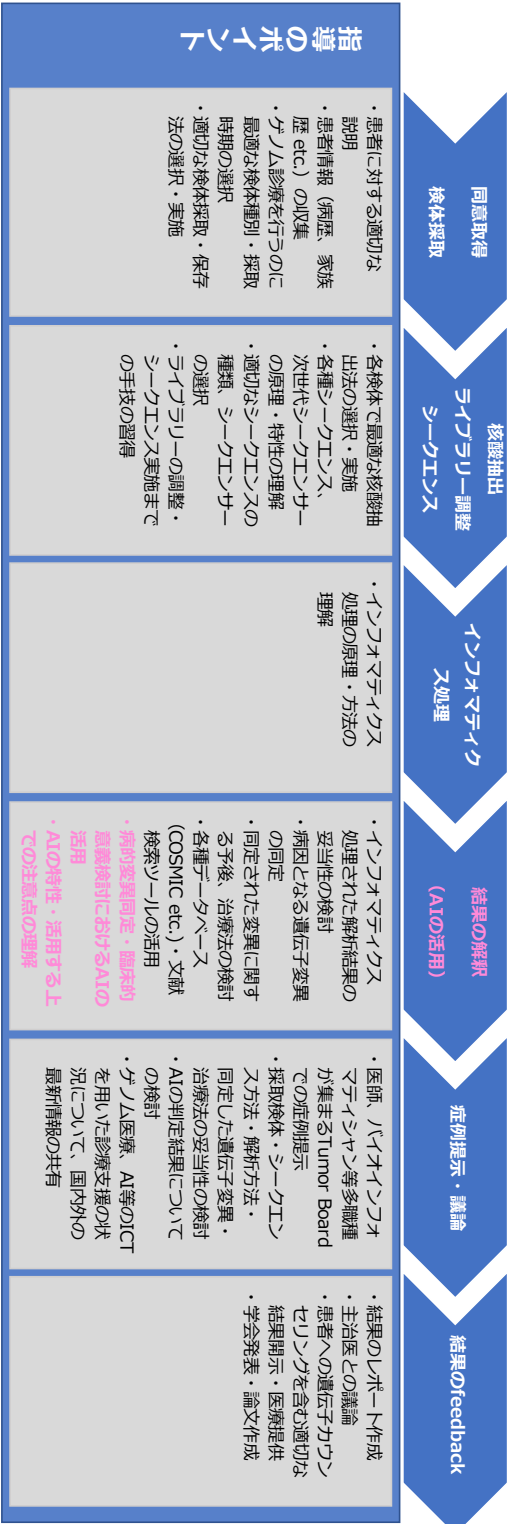
また、OJT を通じた指導の要点から明らかになった専門家育成カリキュラム案を、特に重要と思われる項目の一部(概要版)と、実際に学修する際に必要なキーワードも含め記載した(詳細版)を最後に記載する。

(倫理面への配慮)

本研究でレビューした東大医科研病院でのWfGを用いた血液腫瘍における臨床シーケンス研究は、東大医科研での倫理審査委員会の承認を得て実施されている。また「世界医師会ヘルシンキ宣言」（2008年10月修正）、文部科学省・厚生労働省「人を対象とする医学系研究に関する倫理指針」（平成29年改訂）、「ヒトゲノム・遺伝子解析研究に関する倫理指針」（平成29年改正）を遵守して行われている

AIを活用したゲノム医療のOJT

AIを活用したゲノム医療の流れ



【図 4】

3. 結果

指導事例

1) 症例1

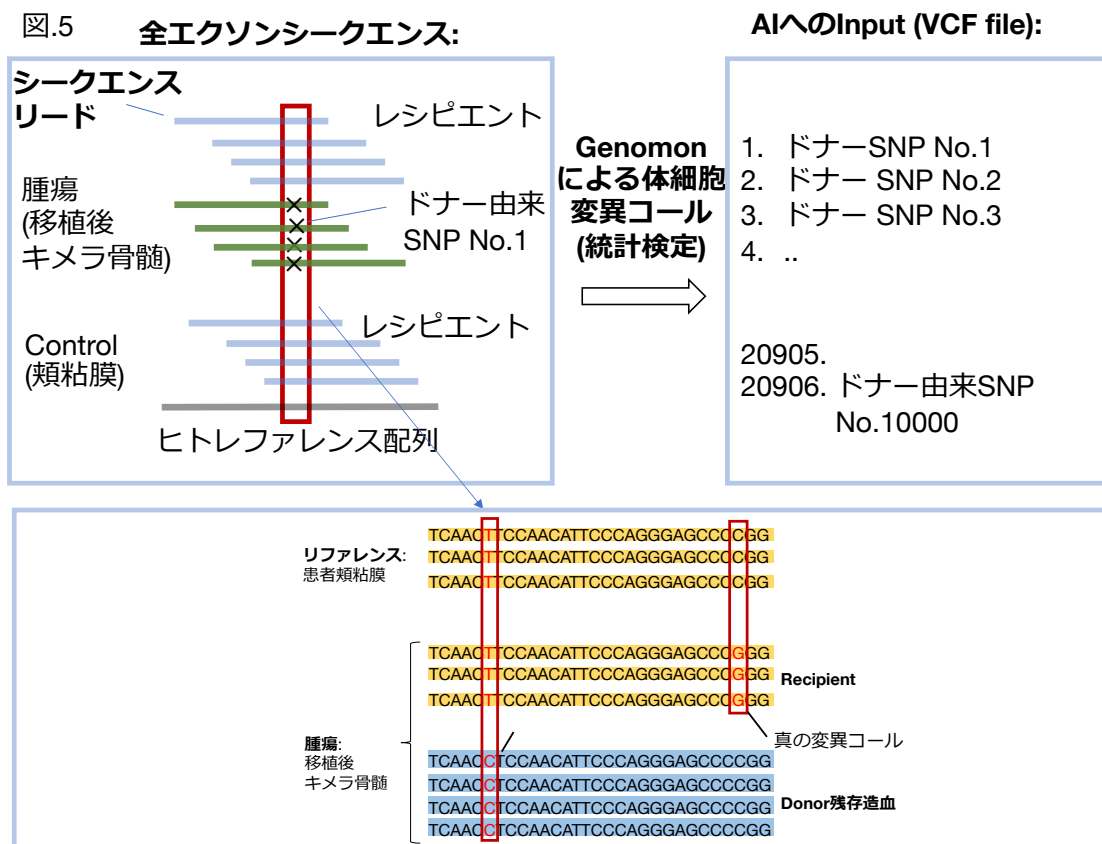
研究協力者1は、急性リンパ性白血病(Acute lymphoblastic leukemia: ALL)の造血幹細胞移植後に再発した患者に対し、再発時骨髓を腫瘍検体、患者頬粘膜を正常対照として全エクソンシーケンス(Whole exome sequencing: WES)を施行した。インフォマティクス処理は、スーパーコンピュータ上に実装されたNGSデータ解析パイプライン、 Genomon (<http://genomon.hgc.jp/exome/index.html>)、を用いて以下の手順で行った。 まずシーケンサーから出力されたリード(FASTQファイル)から、アダプター配列、サンプル固有のインデックス配列などを除去した。その後、シーケンスリードのヒト参照ゲノム配列(hg19)に対する位置を同定(マッピング処理)したデータを生成した(BAMファイル)。次に腫瘍と正常対照のBAMファイルデータを用いて統計的仮説検定を用いた解析を行い、腫瘍検体において、正常対照と比較した場合に有意に変異が観察される位置情報の検出を行った(変異コール)。統計的仮説検定は、Fisherの正

正確率検定を用いて行い、帰無仮説の棄却域は p 値 <0.10 で設定した。変異コールの結果、腫瘍検体においてアミノ酸置換を伴わない変異とアミノ酸置換を伴う変異が合計27,606個、検出された。次に変異コールにより得られた変異とその位置情報のデータ(VCF ファイル)を、AI用のインプットファイルとして出力した。

AIに病名、年齢、性別の情報とともに腫瘍検体の体細胞変異情報として上述のVCFファイルを投入したところ、AIは腫瘍検体において1 Mbあたりの変異塩基数が多い、つまり「体細胞遺伝子変異の数が多い (Tumor Mutation Burden-high: TMB High) 」と判断した。さらに、AIはActionとして、TMB Highの固形がん症例では、変異蓄積により、ネオアンチゲンが腫瘍細胞表面にMHCを介して表出され、CTLの標的となり得ることから免疫チェックポイント阻害薬の効果が高いとする既報等を根拠に(PMID: 25765070)、PD-1阻害薬を提示した。しかし、研究協力者1は、AIの提示した結果が妥当であるのかを評価できなかった。B-ALLではこの様なTMB highの症例は1 %程度にすぎず、極めてまれであることが報告されている事から研究代表者はその結果を疑った(PMID: 28420421)。そこで研究代表者らの指導の元、研究協力者1が注意深く結果を吟味する

と、本例では実際には変異数が多い訳ではなかった。この原因は、同種移植後の患者内には、患者由来とドナー由来の細胞が混在しており、患者口腔粘膜を正常対照として骨髄の遺伝子変異を検出した場合、正常造血を行なっているドナー細胞に観察される生殖細胞変異を患者の体細胞変異として、インフォマティクス処理の過程(変異コール)で検出してしまうためであった(図5)。この症例の解釈には遺伝子変異の情報だけではなく、患者背景とその治療歴、変異コールなどのインフォマティクスの原理を考慮する必要があり、このように多数の因子が関係する結果の解釈は、現状AIのみでの解釈は不可能であり、専門医による解釈が必要である事がわかる。

【図 5】



2) 症例2

研究協力者2は、再発難治性のALL患者に対し、再発時骨髄を腫瘍検体、患者頬粘膜を正常対照として血液腫瘍に関わる51遺伝子のパネルシーケンス(Targeted deep sequencing: TDS)をまず行なった。しかし、本症例においては、TDSでは疾患の原因となるドライバー変異を同定することができなかった。次に全エクソンシーケンス(WES)を施行したが、やはりドライバー変異と思わ

れる変異は検出できず、AIも臨床的に有用と思われる結果は提示しなかった。

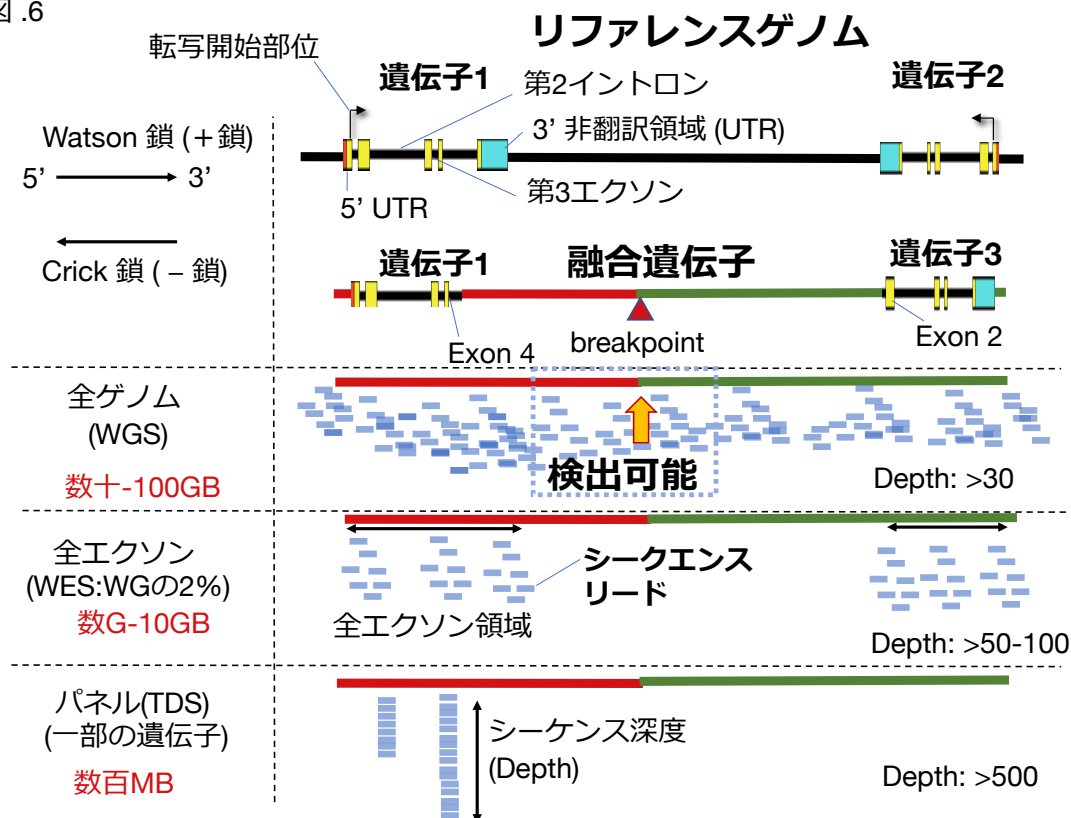
この時点で、研究協力者2は本例では治療標的になり得る変異はないと解釈していた。しかしながら、研究代表者らの指導の元、研究協力者2は、全ゲノムシーケンス(WGS)、及びRNAシーケンス(RNAseq)を再発時骨髄のRNA検体に対して施行したところ、WGSとRNAseq両方の検査において、RCSD1-ABL1 融合遺伝子を検出した。この融合遺伝子が本症例において、難治性である原因の可能性が示唆された。更に、この融合遺伝子に対してImatinib、Dasatinib等のチロシンキナーゼ阻害薬が有効であるという症例報告があり(PubMedID:27125982, 28057740)、AIも同様の結果を提示した。これらの薬剤は慢性骨髄性白血病では既に標準治療として使用されている薬剤であり、実際に投与には至らなかったものの、自費診療の範囲では現実的に治療の選択肢になり得る結果であった。

現在、日本ではがんゲノム医療の全国レベルでの均一化を目的に、主に複数の遺伝子を対象にしたTDSが先進医療として進められ、保険収載されることが決まった。TDSは、より網羅的なシーケンス解析（例：WES/WGS(図6)/RNAseq）と比べた場合に、1) 網羅的なシーケンス解析よりもシーケンスにかか

るコストが低く、費用面で優れている。2)一部のがんにおいてはドライバー変異(特にホットスポット変異)を効率的にスクリーニングでき、1遺伝子あたりのシーケンス量が増えるため、対象となる遺伝子変異に関してはより高感度な変異の検出が可能、という二つの利点を持つ。つまり、TDSは、事前に想定したホットスポット変異などを、腫瘍割合が少ない検体において高感度に検出する場合に限れば有用な検査法と言える。しかし一方で、遺伝子パネル検査では、カバーできるドライバー変異が限られている事を事前に理解しておく必要がある。例えば、ホットスポット変異部位以外のミスセンス変異、比較的稀ながん関連遺伝子、ゲノムの大きな構造異常(融合遺伝子(図6)、重複、挿入、欠失など)、等に関しては対応できない事が多い。特に本症例のように融合遺伝子が原因となる事が多い疾患では、RNAseq、またはWGSを施行しなければ高頻度に見逃しが生じる(図6)。AIは現状ではシーケンス結果を解釈する事しかできないため、どのシーケンス方法を選択するかは専門医が決定する必要がある。制限されたシーケンス法で得た結果のみを解釈してしまうと、治療標的になり得る変異を見逃してしまう危険性があるため、疾患背景を理解し、各患者に適切なシーケンス方法、またそれに合わせた核酸抽出法を選択する必要がある。

【図 6】

図 .6



3) 症例3

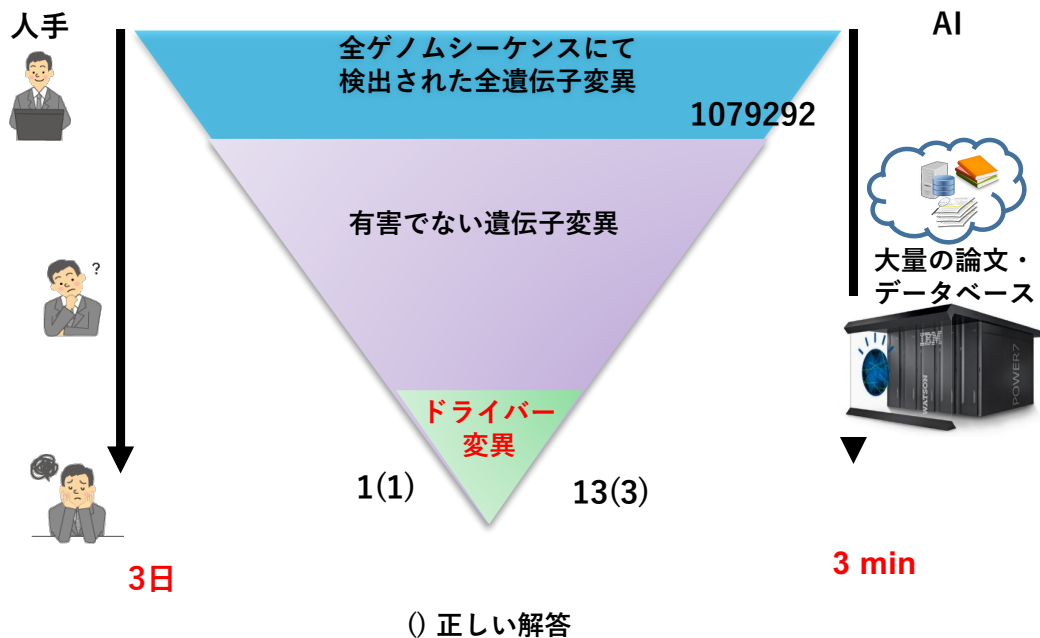
研究協力者3は、ALL移植後再発患者の骨髄検体に対して、最初からWGSを施行した。本症例では、正常対照として患者頬粘膜が入手できなかった。研究代表者らが、研究協力者3に対して最初からWGSを施行した理由を問うと、症例3は症例2と同じ疾患であり、症例2では融合遺伝子がドライバー変異であったとい

う経験を事前に研究協力者2と共有していた為ということであった。

本症例のインフォマティクス処理は症例1と同様の手順で行ったが、変異コールの際には正常対照の代替検体として他の患者頬粘膜DNAを用いた。また、Genomonパイプラインに含まれる構造変異検出ツール(Genomon-SV)を用いた構造変異解析も本症例では行ったが、その際のフィルタリングの閾値パラメータはデフォルト値を用いた。

Genomonでの変異コールの結果、1,032,522の遺伝子変異と46,770の構造異常が検出された。まず研究協力者3が結果を解釈したところ、約3日間を要し、1つの構造異常（12p欠失:12番染色体短腕欠失）を同定した。次にAIでシーケンス結果の解釈を試みたところ、AIは結果を数分で解釈し、合計13個の異常をドライバー変異として検出した。この中には研究協力者3が同定した構造異常に加え、新たに1つの変異（SETD2）と2つの構造異常（13q欠失:13番染色体長腕欠失、17q欠失）が含まれていた。最後に研究協力者3が、研究代表者らの指導の元でAIの結果を約2時間かけて解釈したところ、実際にこれらの3つは疾患の病態形成に関連していると考えられた(図7)。

【図 7】



本症例では下記に記す手順 1 でシーケンス結果の解釈を行ったが、特にデータ量が膨大であった事から、最初のSTEP1に3日を要した。この事は、迅速にシーケンス結果を臨床医に返却する事が求められるゲノム医療の実践において特にボトルネックになり得ると考えられた。専門医による解釈というボトルネックを克服し、解釈時間を短縮するためには、下記に記す手順2の様に、ステップ1を省略し、最初からAIを使う手順に変更するべきであると考えられた。たまたま本症例の場合は、専門医の推論がAIの推論結果に含まれている為、手順2でも結果の見逃しは生じない(図8)。一方、専門医の評価のみで判明する場

合もあり得る(図9)。そこで、見逃しをできるだけ少なくし、かつ専門医による解釈を効率的にするには、例えば手順3の様に、複数のAIを用いてあらかじめ評価した結果を専門医が評価する手順も有用ではないかと考えられた(図10)。

手順1:

(STEP1) 専門医による結果の解釈ステップ(3日間)

(STEP2) AIによる結果の解釈ステップ(数分)

(STEP3) STEP1,2で得た専門医とAIの結果の比較と評価を行うステップ(2時間)

手順2:

(STEP2) AIによる結果の解釈ステップ(数分)

(STEP3') STEP2で得られたAIの結果を専門医が評価するステップ(2時間)

手順3:

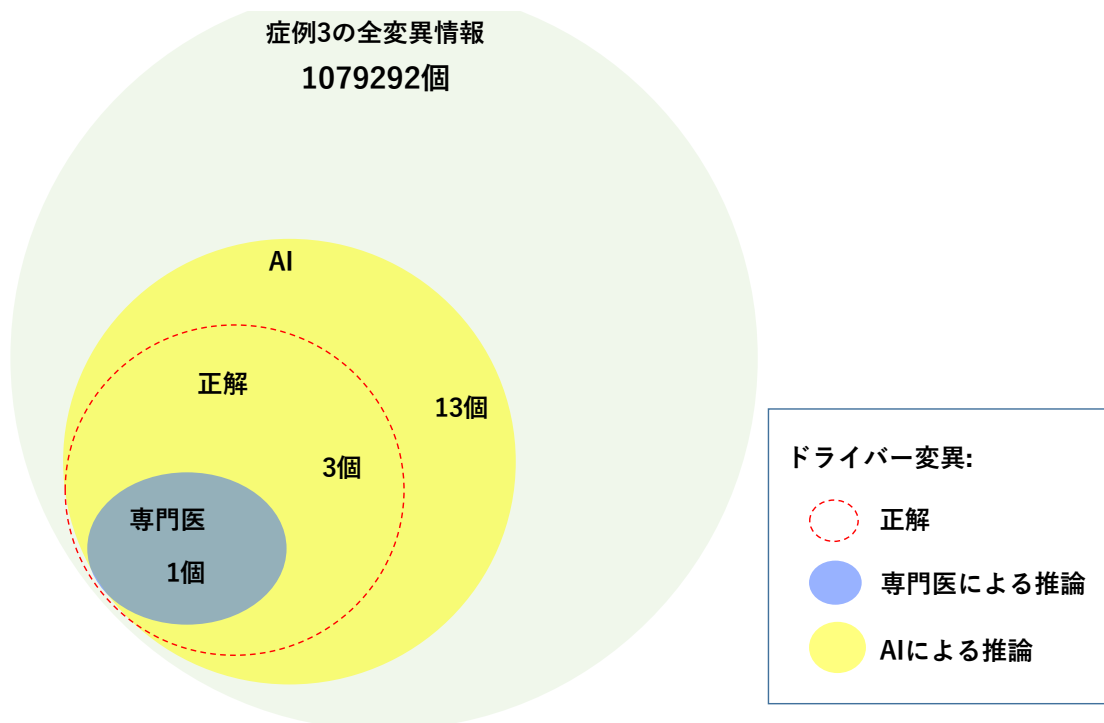
(STEP2) AI-No.1による結果の解釈ステップ(数分)

(STEP2') AI-No.2による結果の解釈ステップ (数分)

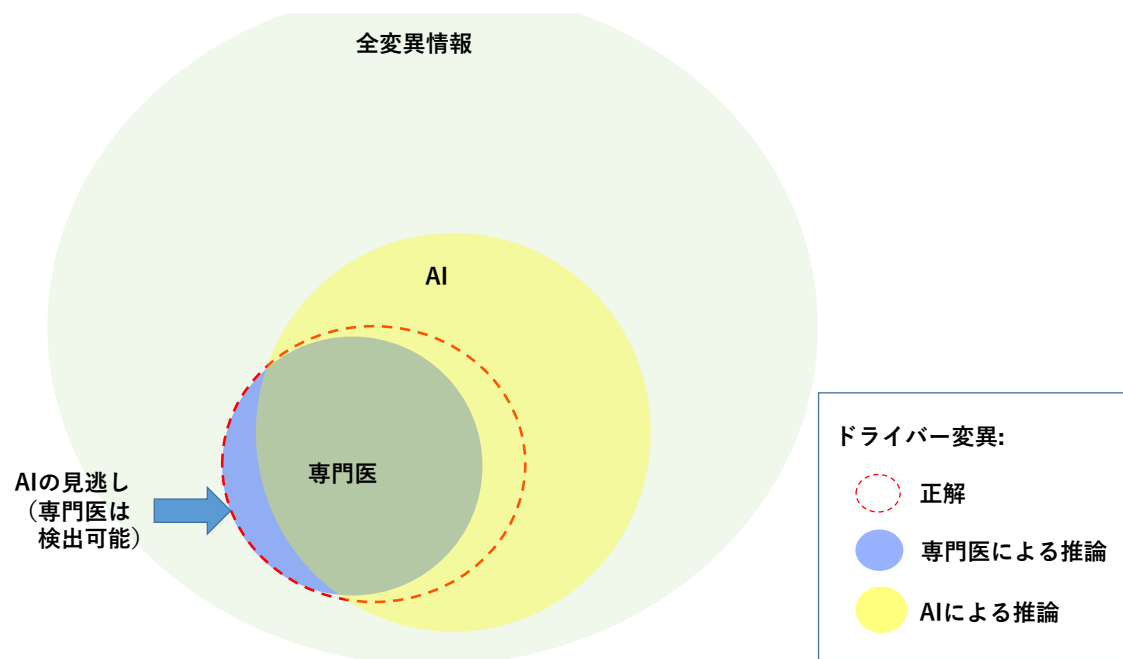
(STEP2'') AI-No.3による結果の解釈ステップ (数分)

(STEP3') STEP2-2''で得られたAIの結果を専門医が評価するステップ(2時間)

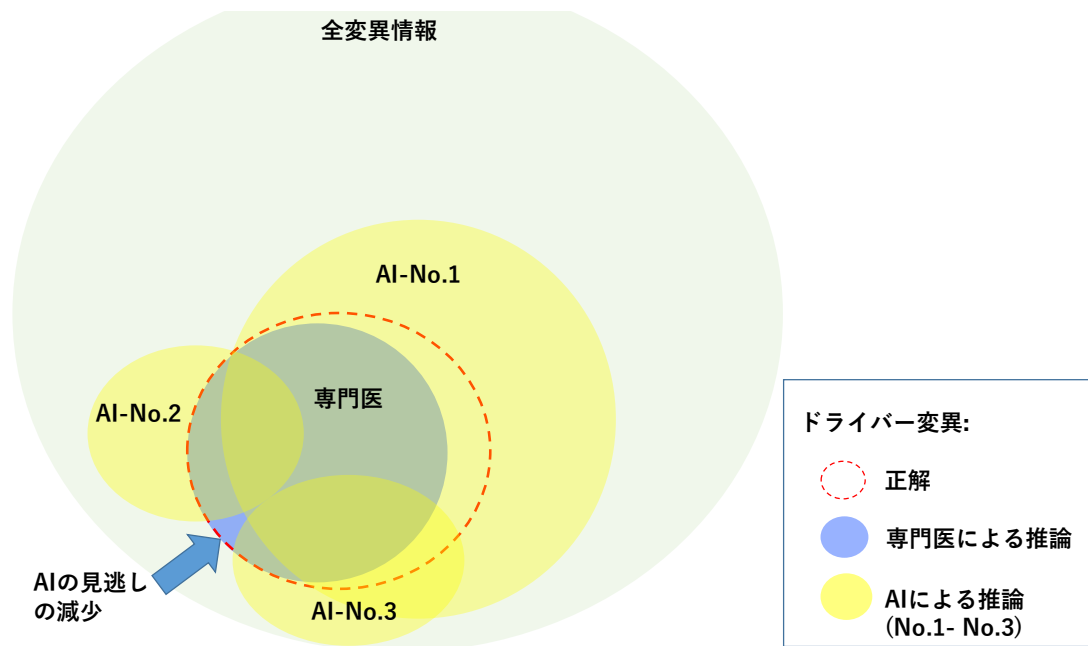
【図 8】



【図 9】



【図 10】



本症例の様にシーケンス結果が膨大である場合、人間の眼のみで結果を確認し、臨床的な解釈をするには多大な労力と時間を要し、重要な変異を見落とす危険性も高くなる。対してAIは短時間で大量のデータを解析する事が可能なため、人間だけでは数日かかる過程を数分で行う事ができる。この特性を生かして複数のAIを組み合わせれば誤ってデータを見落とす心配も少なくなると考えられる。今後ゲノム医療の進歩に伴い解析するシーケンスデータが増大する事が予測される中、単独、あるいは複数のAIを専門医の解釈と組み合わせ有効

に活用する事により、短時間でシーケンス結果を解析でき、診断仮説の形成や治療方針立案において効率的に情報を得る事が可能となると考えられた。

4) 症例4.

研究協力者 1 は、化学療法後に再発した ALL において、骨髓を腫瘍検体、患者頬粘膜を正常対照として WES を施行した。変異コールの結果、816 個の変異が検出された。このうち重要なドライバー変異候補として、AI と研究協力者 1 は EZH2 変異を推論した。次に WGS を施行したところ、828,054 変異が検出された。研究協力者 1 は EZH2 変異の他に、白血病の病態形成に関わる極めて稀だが重要なゲノム構造変異として、TAF15-ZNF384 融合遺伝子(PubMed ID: 21504714、21962896)を検出した。一方、同様の変異はこれまで ALL において十数例程度の症例報告しかなく、その希少性故に代表的ながんのデータベースにはゲノム構造変異として、解析した当時において登録がなかった。そこで AI は TAF15-ZNF384 融合遺伝子を重要な遺伝子変異とは推論しなかった。しかし、AI の開発 Group にこの事を feedback し学習すると、ドライバー変異候補として call されるようになった。

本症例でも、症例 2 と同様の事が考察された。例えば、融合遺伝子等のゲノムの大きな構造変異(図6)、ホットスポット変異部位以外のミスセンス変異、比較的稀ながん関連遺伝子、等に関してはTDS(図6)では対応できない事が多い為、融合遺伝子が病態形成において重要である事が多い ALLの様な疾患では、RNA seqか、本例の様にWGSを施行しなければ高頻度に見逃しが生じると考えられた。また、特に本例の様に、癌のデータベースに登録がない、症例報告や症例シリーズ報告 (case series) レベルの根拠しかない変異に関しては、ドライバー変異としての学習が出来ていない為に、AIが重要な変異として推論しない可能性がある事にも留意が必要であると考えられた。この例の様にAIの出力を鵜呑みにするのではなく、人間が最終的にAIの出力結果を確認する事が重要である。

5) 症例5.

研究協力者2は、ALLに対して化学療法後に骨髓異形成症候群(MDS)を発症した男性において、MDS診断時の骨髓を腫瘍検体、患者頬粘膜を正常対照としてWESを施行した。変異コールの結果、腫瘍検体において体細胞変異が1,725個、正常対象において、リファレンスゲノムと異なる生殖細胞変異候補として32,7

66個が検出された。専門医がまず体細胞変異の結果を解釈したところ、約3日間を要し、2つの変異(PTPN11遺伝子のホットスポット変異 (p.E69K)、ETV6 のフレームシフト変異)を同定した。次にAIでシークエンス結果の解釈を試みたところ、AIは結果を3分で解釈し、合計8個の変異をドライバー変異候補として検出した。この中には、専門医がドライバー変異候補として推論したPTPN11変異は含まれていたものの、ETV6変異は含まれていなかった。AIは残り7個の変異は臨床的意義不明の変異(variants of unknown significance: VUS)と推論しており、この結果は専門医の解釈と一致した。一方、専門医のみが推論したETV6に関しては、フレームシフト変異により、ETV6のDNA結合ドメインであるETSドメインを欠失した機能喪失が起これと考えられた。MDSやAML の体細胞変異の遺伝学的背景において機能を喪失したETV6変異のエビデンスが豊富にある(PMID: 28835720、PMID: 24220272)事を踏まえ、最終的に ETV6とPTPN11がMDSの病態形成に関連していると結論づけた。

次に、正常対照として頬粘膜を解析して得られた32,766個の生殖細胞変異の解釈において、専門医は解釈に2週間以上を要し、重要な変異はないと推論した。一方、AIを用いてシークエンス結果を腫瘍検体と想定した解析を行ったと

ころ、わずか3分で結果を推論し、22個の変異を重要な変異としてlist upした(図11)。AIの解析結果を専門医が詳しく検討したところ、PALB2の機能喪失生殖系列変異が含まれている事がわかり、実際に本症例ではサンガーシーケンシングによる確認検査で変異が確認された。PALB2は、BRCA2の「パートナー」と呼ばれ、BRCA2タンパクと相互作用しDNA修復に関与する遺伝子である。PALB2変異は、常染色体優性遺伝形式をとり、乳がん等のリスク増加とも相関すると報告されている(PMID: 25099575)。さらに、PALB2変異は治療関連骨髄腫瘍(t-myeloid neoplasms : t-MN)の発生リスクを高める可能性が示唆されている(PMID: 28835720)。以上の事から、専門医は、本例はt-MNの発生素因であるPALB2変異に、殺細胞性抗がん剤治療が加わり、t-MNを発症した可能性が高いと推論した。本例では、元々この様な偶発的所見の開示希望があった。そこで、患者や家族の健康管理に役立てていくために、当院の遺伝性腫瘍を専門とする臨床遺伝専門医らと連携し、当院での遺伝カウンセリングを含めた今後の対応を協議した後に、患者と家族に対して遺伝カウンセリングを実施して結果を説明する事となった。

本症例5や症例3では、人間の解釈可能なデータ量には制限があるが、AIにはそ

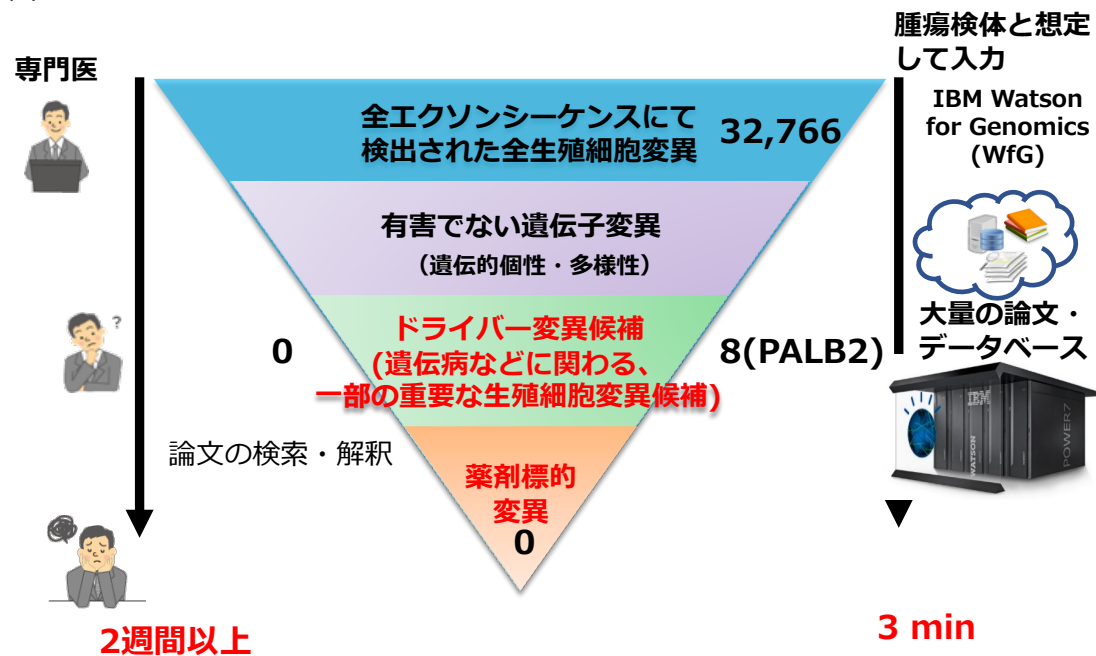
れが無いという特徴がよくわかる。つまり、解析データ量が膨大になればなるほど、AIを活用する事で、人間による見逃しを防ぎ、臨床シーケンスにおける解釈の精度を高める事に繋がると考えられる。実際、研究代表者らは、当院におけるAMLなどの血液がん症例(n=186)での検討において、Single nucleotide variant: SNVの解釈に関して、AIと専門医の一致率は実用レベル(81%)である事を確認している。一方、症例4では融合遺伝子、本症例5では体細胞変異のETV6の見逃しが、AIにあった。この原因は、TAF15-ZNF384 融合遺伝子 やETV6が、それぞれB-ALLやMDSやAML の体細胞変異の遺伝学的背景において重要な変異として、紐付けされていない、つまりAIが関連を未学習であり、学習済みデータベースに無い事に起因すると考えられる。この様にAIの出力結果にも見逃しや誤りがあり得る事から、最終的な判断を担うのは医師であることにも留意すべきである(出典：平成29年度厚生労働行政推進調査事業補助金 「AI等のICTを用いた診療支援に関する研究」 報告書 [研究代表者：横山和明]、平成30年12月9日 医政医発 1219 第1号 「人工知能(AI)を用いた診断,治療等の支援を行うプログラムの利用と医師法第17条の規定との関係について」

Available from: www.mhlw.go.jp/content/10601000/000468150.pdf。

また、本例の様に腫瘍組織と正常対象組織のNGS解析に伴い、偶発的所見として生殖細胞系列変異が認められる場合がある事にも留意すべきである。従って、研究の説明を行う前に、本人への結果の開示も含めた対応について本人と相談し決めておく事、また偶発的所見があった場合に、本人や家族への遺伝カウンセリングが可能な体制の事前構築も必要であると考えられる(出典：ゲノム医療実用化推進研究事業「メディカル・ゲノムセンター等における ゲノム医療実施体制の構築と人材育成に関する研究」サブテーマ2 別冊報告書「偶発的所見・二次的所見への対応についての検討と提言」<https://www.mhlw.go.jp/content/10901000/000486817.pdf>)。

【図 11】

図.11



6) 症例6.

研究協力者 3 は、1 年前から検診で指摘される好中球著増を主訴に、精査目的で当院に紹介受診した、60 歳の骨髄増殖性疾患症例の NGS を担当する事となった。まず、研究協力者 3 は、事前に鑑別診断を考える事にした。当該症例の骨髄検査結果を確認したところ、骨髄では有核細胞数が著増しており、著明な過形成髄の所見であった。骨髄における細胞増加の主体は成熟した好中球

であり、芽球の増加は認めなかった為、急性骨髄性白血病は否定され、骨髄増殖性疾患に合致する所見と考えられた。一方で、骨髄では、MDS に特徴的な血球の異形性が散見され、MDS としても診断は矛盾しない所見だと考えた。

鑑別疾患として、成熟した骨髄球系が増加する慢性骨髄性白血病（CML）や分類不能型 MDS が挙げられた。染色体検査は正常であり、BCR-ABL-FISH 法も陰性である事から、CML は否定できると考えた。以上の臨床検査結果や病歴を踏まえ、研究協力者 3 は、好中球の腫瘍性増殖を示す病態、慢性好中球性白血病（CNL）の可能性が高いと判断した。次に研究協力者 3 は、PUBMED で、CNL に特徴的な遺伝子異常を文献検索したところ、CNL ではコロニー刺激因子 3 受容体（CSF3R）の活性化変異が遺伝学的特徴としてこれまで報告されており、当該変異は、WHO の CNL での重要な診断基準項目の一つとして挙げられている事が分かった(PMID: 23656643)。次に、研究協力者 3 は、パネル検査で CNL に特徴的な CSF3R 変異が検査可能か調べる事にした。試薬会社から公表されている、パネル検査にてカバーされるゲノム領域の list として、Browser Extensible Data (BED)形式 file を WEB page から入手した。BED file には、カバーされるゲノム領域の染色体（ゲノム）上の位置情報、具体的に

は染色体の番号、プライマー配列の開始位置、プライマー配列の終了位置が書かれている。研究協力者 3 が、UCSC Genome Browser を用いて、パネルの PCR 増幅産物がカバーする、CSF3R の当該領域の座標を検索したところ、CNL のホットスポット変異である CSF3R の膜近傍領域の p.T618I 変異はカバーされるが、CNL や類縁疾患の非定形 CML の約 2、3 割で検出される細胞内ドメインの変異は、当該パネル検査ではカバーされない事が分かった。そこで、研究協力者 3 は、CSF3R の細胞内ドメインも検出する必要があると考え、パネルではなく、全エクソンシーケンスを解析手法として選択する事にした。検体は、患者骨髓を腫瘍検体、患者頬粘膜を正常対照として用いた。体細胞変異コールの結果、1175 個の変異が検出された。研究協力者 3 はコロニー刺激因子 3 受容体 (CSF3R) の p.T618I 活性化変異, ASXL1, U2AF1 の変異が (PMID: 23656643、PMID: 31366621)、本例において病態形成に寄与した重要な変異として推論し、特に、CSF3R の活性化変異が CNL に疾患特異性が高いと結論づけた。

以上の結果から、研究協力者 3 は、本例は CNL とである診断されると結論づけた。さらに、本症例の様に CSF3R の膜貫通領域近傍のホットスポット変異

の場合、下流の JAK 2 キナーゼシグナル伝達の活性化が起こる為、JAK1/2 阻害剤である Ruxolitinib の効果が報告されている (PMID: 23656643) 。そこで研究協力者 3 は、治療標的として Ruxolitinib を推論した。一方、AI は CSF3R を治療標的として推論しなかった。にもかかわらず、CSF3R 変異に関する AI の出力情報の欄を見ると、“この hotspot 変異に対しては Ruxolitinib の有効性が報告されている”と action の記載があった。このことを研究協力者 3 が、AI の開発 Group に feedback すると、直ちに actionable 変異として Ruxolitinib が call されるようになった (図 12) 。

本症例では、AI はドライバー変異として CSF3R を推論し、それに関連した文献記述情報として、Ruxolitinib の有効性を含んだ記述を出力可能であった。にも関わらず、AI は Ruxolitinib を CSF3R と関連した action としては推論しなかった。しかし、開発チームに Feedback 後には、AI は CSF3R と Ruxolitinib の薬剤標的としての関連を学習したことから、action として出力する事が可能であった(図 12)。

研究協力者 3 は、以下の様にこの事例における AI の解析ロジックを解釈した。AI は骨髄増殖性疾患における重要なドライバー変異として、CSF3R を事

前に学習して情報を格納していた。また、CSF3R に紐づいた関連記述として Ruxolitinib の有効性を含む文献記述も同時に格納していた。しかし、AI は薬剤標的としては、関連を未学習であり、Ruxolitinib を薬剤標的として出力する事はできなかった。しかし、学習（開発チームへの feedback）により、CSF3R と Ruxolitinib の関連を紐付ける事が可能となり、推論精度が向上した。

この指導例の様に、AIは学習により結果の出力が変わり得る事、具体的には推論精度が上がる事があり得る事、も理解する必要があると思われた。また、文献の記述内容やそのロジック、質問の内容自体を理解しているわけではない事に、特に留意すべきと思われた。つまりAIが結果を紐付けて出力する為には、事前に関連を学習する必要がある。学習により推論精度が本例では向上したが、逆に精度が低下する事例もあった。AIはあくまでも膨大な事前学習データに基づき、質問に対する単純な相関を元にランク付けして結果を出力する便利な検索ツールと捉えるべきだと考えられた。質問の内容自体を理解する事や、さらに高度な、出力した結果から原因や関係を見つけ出すなどの高度な推論は人間には当たり前で、AIには難しいので人間が担うべき仕事として当面は残ると考えら

れる。以上より、AIを使いこなす上で、AIの特性、inputに対してoutputを出力する仕組み(=推論のロジック)、特にAIの基盤技術である、機械学習の仕組みを理解しておく事、最終的な結果を人間が確認する事が重要であると考えられた。

【図 12】

図 .12

学習前 AI output:

1	Actionable Alterations	0	FDA Approved for Myeloproliferative Neoplasm	0	Therapies with Clinical Trial(s)	1	FDA Approved for Other Indication(s)
CSF3R	<i>T618I</i>	-	-	-	-	-	-

学習後（開発チームにJAK 2 阻害剤を actionとしてfeedback）

1	Actionable Alterations	0	FDA Approved for Myeloproliferative Neoplasm	0	Therapies with Clinical Trial(s)	1	FDA Approved for Other Indication(s)
CSF3R	<i>T618I</i>	-	-	Phase 2 Trial	-	-	-
				4 Ruxolitinib			

7) 症例7.

研究協力者 1 は、高リスク MDS を発症した 70 台男性において、MDS 診断時の骨髄を腫瘍検体、患者頬粘膜を正常対照として WES を施行した。変異コールの結果、腫瘍検体において体細胞変異が 1,385 個検出された。専門医と

AI は共に、本症例における病態形成に寄与した変異として TP53 のミスセンス変異(p.R273C、p.G244S)が両アレルに起こった事によると推定した。実際、TP53 変異を持つ急性骨髄性白血病(AML)や MDS は極めて予後不良であり、抗癌剤を含めた治療に抵抗性である事が報告されている (PMID: 28177873)。

専門医は、可能性のある action として、脱メチル化剤を以下の根拠で推論した。1)脱メチル化剤の一つ、アザシチジン(Azacitidine)は、本邦において、すでに高リスク MDS において承認されている、2)TP53 変異がある MDS や AML では、比較的変異を持たない例に比べて、その機序やパスウェイは不明だが、脱メチル化剤(デシタビン)が奏功しやすく、治療を受けた例の予後を、TP53 変異を持たない例と比べると、同等の生存率までその予後が改善する可能性がある事(PMID: 27959731)、であった。一方、AI は、パスウェイ Target ではない事から、脱メチル化剤を action としては提示しなかった。研究協力者 1 が、解析結果を主治医に feedback したところ、本例では、実際に MDS の治療として、抗がん剤による化学療法ではなく、直ちにアザシチジンによる治療が選択された。患者は 1 コースの治療で、芽球が著明に減少し、血小板数の著増などの臨床効果が得られ輸血依存から脱した。3 コース目の効果

判定では、完全寛解を達成した。研究協力者 1 は、開発チームにこの結果を数回 feedback したが、開発チームが AI に対してこの結果を学習させる事はなかった。その理由は、AI は薬剤に関連するパスウェイと紐付けて action を推論し、出力するが、アザシチジンは TP53 のパスウェイターゲットではない為、action として紐付ける学習を AI にさせる事はできない、という事であった。この様に、文献で仮に有効性が示されていても、action が直接のパスウェイターゲットでない場合等には、AI の推論ロジックからは外れる事があり、学習が行われず、有効な action が推論されない事もあると考えられ、この点に留意する必要があると考えられた。また本例においても、最終的な結果は人間が確認する事が重要であると考えられた。

8) 症例 8.

研究協力者 2 は、高リスク MDS から AML を 2 次性に発症した 70 台男性において、診断時の骨髓を腫瘍検体、患者頬粘膜を正常対照として WES を施行した。腫瘍検体において体細胞変異が 887 個検出された。専門医と AI は共に、本症例における病態形成に寄与した変異として NRAS のホットスポットの

ミスセンス変異(p.G12A)を、本症例における病態形成に寄与した変異として推論した(PMID: 28835720、PMID: 24220272)。一方で AI は、本変異は、AML において予後良好の変異だと推論し、その根拠となった文献を出力した(PMID: 26980726)。研究協力者 2 はこの推論に疑問を持ち、実際にその文献を読んで解釈したところ、実際には、特定の予後良好な遺伝子異常を持つ AML の(t(8;21) or inv(16))において、さらに KIT や F L T 3 変異を持たない例に限れば予後良好という記載であり、本例の様に MDS から AML に進展した 2 次性白血病や AML 全般において当てはまる訳ではなかった。研究協力者 1 が、開発チームとこの推論結果の誤りについて議論したが、AI の推論ロジックの詳細は公開されなかった。しかし、議論を通じて研究協力者 2 と代表者は、推論結果の誤りは、AI が学習したコーパス（知識単位）に問題があった可能性があった為であると推論した。コーパスとは AI が自然言語処理を行ったテキストをベースに、人間が内容を吟味した上で、AI に入力する知識単位の事を指す。具体的には、本例では、当該部分が記載された構文が、二重の主語、連体修飾を含み、複雑な構文であったことから、単純なコーパスとして述語構造を収集するだけでは、誤りが起きる可能性があると考えられた。本文献では、

「t(8;21)陽性で、KIT 変異を持たない、NRAS 変異陽性の AML 例は予後が

良い」という記述されている。この場合、単純な自然言語処理による構文解析

では、“NRAS 変異陽性の AML 例は予後が良い”という部分だけ抜き出されて

コーパスとして情報収集される可能性がある。この場合、AI に学習させる前に

人間が、収集した情報の確認を行う手順となっているが、もし人間がコーパス

として入力する情報を確認する段階で誤りを見逃した場合には、そのまま AI

の学習に用いられ、本例での推論の様に、NRAS 変異陽性の AML は予後が良

いと AI が推論間違いをする可能性があると考えられる。本例においては、

1)t(8;21)陽性で、2)KIT 変異を持たない、という AML に係る連体修飾の語句

の関連を、自然言語処理による情報の収集段階の過程で、適切に構文処理され

ずその点が考慮されなかった事が誤りの原因と考えられる。このように、複雑

な構文構造を持つ自然言語テキストの場合、AI がそのまま知識源として使いこ

なすのは現状ではまだ難しい可能性があると考えられた。

4. 考察

一連のAIを活用したゲノム医療のOJTを実際に研究協力者として3人の大学院生（医師）が経験した中で、昨年度報告した事例も含めて、代表的な指導事例8例とその指導ポイントを提示した。いずれの指導ポイントも、医学部卒業後数年程度の大学院生3名に対して、何故自分たちが結果の解釈につまずいたのかを理解させる点において昨年度に引き続き有用であった。

症例1では、研究協力者1は、腫瘍検体のシーケンス結果において多数の変異がコールされ、AIが腫瘍細胞において変異が多いと判断してしまった理由に当初気づく事ができなかった。この理由を解釈するには、患者背景や治療歴、インフォマティクス処理の原理を事前に理解しておく必要があった。この点は、自身の知識を強化することと、インフォマティクスを専門とする研究者との連携を強化するという対策を講ずるべきであろう。

次に、症例2においては、患者疾患がALLであり、融合遺伝子異常が原因となる事が多いことから、最初にシーケンス方法を選択する段階でRNAシーケンスや全ゲノムシーケンスを考慮する必要があった。実際、本症例では融合遺伝子がドライバー変異であり、最初からこれらのシーケンスを行っていたら、より早い段階で原因遺伝子異常を同定し得た。この症例に適切なシーケンス法

を選択するには、対象とする疾患の遺伝学的背景や各種シーケンス法の欠点とその長所の理解が必要であった。更にこの症例では治療選択肢になり得る薬剤があり、研究協力者は、適切でないシーケンス方法を選択する事は、治療標的になり得る変異の見逃し、また治療の遅れに直結するという危険性を実感した。

次に症例3においては、研究協力者3は膨大な変異の解釈をする際に、人間の解析量には限界があり、一方でAIにはそれがなく効率的に有益な情報を得る事ができるという事を実感した。さらに本症例ではAIを併用する事により人間が見落としていた重要な変異を検出し得たが、逆にAIの結果のみに頼ると重要な変異を見落とす可能性もあるため、人間の誤りや限界を補うためには、複数のAIを使用する等、特性を理解した上で両者を適切に組み合わせる事が重要である事も理解できた。

次に症例4においては、研究協力者1は、症例2と同様に、臨床シーケンスにおける、適切なシーケンス法を選択する重要性を実感する事ができた。また、AIが推論する場合に、ドライバースコアを計算する根拠となるデータベースに登録がない様な、まれな変異に関しては、ドライバー変異として学習ができてい

ない可能性があり、候補として出力されず、見逃す可能性がある事、最終的には人間がAIの出力結果を確認する事が重要であるという事を実感できた。

次に、症例5においては、症例3と同様に、人間の解析量には限界があり、一方でAIにはそれがなく効率的に有益な情報を得る事ができる、時に人間の見逃しを補完し得る事を実感できた。つまり、研究協力者2は、解析データ量が膨大になればなるほど、AIを活用する事で、人間による見逃しを防ぎ、臨床シークエンスにおける解釈の精度を高める事に繋がる事を実感できた。また、正常対照組織の解析に伴い、偶発的所見として生殖細胞系列変異が認められる場合がある事にも留意すべきである事、そしてその事に関して事前に患者に体制を構築しておく事の重要性や、偶発的所見が見つかった場合の対応を事前に患者と相談しておく必要性も実感できた。

次に、症例6においては、研究協力者3は、事前に病歴や患者背景から鑑別診断を考えて、適切なシークエンス方法を選択する、という一連の流れを実際に実践できた。また、AIの出力結果を鵜呑みにせず、必要なactionを推論できた。また本例の経験を通して、AIは文献の記述内容やそのロジック、質問の内容自体を理解しているわけではない事、を実感できた。さらに、AIは学習により、結果の

出力が変わり得る事、具体的には推論精度が上がる事があり得る事、を理解する事も重要である事を実感した。以上より、AIを使いこなす上で、AIの特性、inputに対してoutputを出力する仕組み(=推論ロジック)、特にAIの基盤技術である、機械学習の仕組みを理解しておく事の重要性を理解できた。

次に症例7においては、研究協力者1は、現状のゲノム解析特化型AIが推論しない事、あるいは推論するのが苦手な事として、変異と直接関係のある直接的なパスウェイターゲットでない場合、あるいは、パスウェイターゲットが明らかとなっておらずデータベースに登録がない場合には、文献で有効性が示される薬剤の記述があっても、癌や薬剤のデータベースに重み付けを行って推論する、ゲノム解析特化型AIの推論ロジックでは、actionとしてみなされない為に見逃しもあり得る、という事を実感できた。本症例においても、AIを使いこなす上で、AIの特性、inputに対してoutputを出力する仕組み(=推論のロジック)、特にAIの基盤技術である、機械学習の仕組みを理解しておく事、最終的な結果は人間が確認する事が重要であると考えられた。

最後に症例8においては、研究協力者2は、特にAIが入力情報として学習したコーパスの素となった文献が、複雑な構文構造を持つ場合においては、自然言語処

理技術の問題から、AIの推論内容自体に誤りがおき得る事を実感できた。AIを使いこなす上で、AIが学習するコーパスデータの基盤技術となる、自然言語処理の概要を理解する事、AIが推論した内容に関して、最終的に人間も結果を確認する事が重要である事を実感できた。

これらの事例を通じて判明した具体的な指導における重要ポイントを以下にまとめておく。

症例1の指導ポイント.

- 1.AIが考慮しない因子を理解させる。例) 患者背景(治療歴)
- 2.AIが関与しないプロセス、特にインフォマティクスの原理について理解させ、インフォマティシャンとの連携を強化する。

症例2の指導ポイント.

- 1.AIが考慮しない因子を理解させる。例) 患者背景(疾患の遺伝学的背景)
- 2.AIが関与しないプロセス、疾患の遺伝学的背景に応じた適切なシーケンス法について、その手法の特徴と共に理解させる。例) 急性リンパ性白血病と

融合遺伝子

症例3.の指導ポイント.

- 1.AIの特性を人間と対比させて理解させる。
- 2.膨大なデータを解釈するためには、AIと人間を適切に組み合わせる事が必要である事を理解させる。例) WES、WGSで出力される変異データの解釈においてAI と人間を組み合わせる有効性。

症例4.の指導ポイント.

1. AIが考慮しない因子を理解させる。例) 患者背景(疾患の遺伝学的背景)
- 2.AIが関与しないプロセス、疾患の遺伝学的背景に応じた適切なシーケンス法について、その手法の特徴と共に理解させる。例) 急性リンパ性白血病と

融合遺伝子

3. AIを使いこなす上で、特にAIの推論における基盤技術である、機械学習の仕組みを理解する。例: inputに対してoutputを出力する仕組み
4. AIの推論結果には、機械学習の特性上の問題として見逃しがあり得る事を、

その見逃しがちなinput情報の属性などと共に理解させる。

症例5.の指導ポイント.

1. 膨大なデータを解釈するためには、AIと人間を適切に組み合わせる事が必要である事を理解させる
2. AIの出力結果には、見逃しがあり得る事を、AIが見逃しがちなinput情報の属性などと共に理解させる。
3. AIを使いこなす上で、特にAIの推論における基盤技術である、機械学習の仕組みを理解する。例: inputに対してoutputを出力する仕組み
4. AIの推論結果には、機械学習の特性上は見逃しがあり得る事を、見逃しがちなinput情報の属性などと共に理解させる。
5. 特に正常対照解析に伴い、偶発的所見として生殖細胞系列変異が認められる場合がある事を、その際の対応と共に理解させる。

症例6.の指導ポイント.

1. AIを使いこなす上で、特にAIの推論における基盤技術である、機械学習の仕

組みを理解する。例：inputに対してoutputを出力する仕組み

2. AIの推論結果には、機械学習の特性上は見逃しがあり得る事を、見逃しがちなinput情報の属性などと共に理解させる。
3. 1、2に関連して特に、AIは学習により結果の出力が変わり得る事を理解する。
4. 1、2に関連して特に、AIは質問の内容自体を理解しているわけではない事を理解させる。
5. AIが推論した内容に関して最終的には人間も結果を確認する事が重要である事を理解する。

症例7.の指導ポイント.

1. AIを使いこなす上で、特にAIの推論における基盤技術である、機械学習の仕組みを理解する。例：inputに対してoutputを出力する仕組み
2. AIの推論結果には、機械学習の特性上、有効なaction等を見逃しがあり得る事を、見逃しがちなinput情報の属性などと共に理解させる。
3. AIが推論した内容に関して最終的には人間も結果を確認する事が重要であ

る事を理解する。

症例8.の指導ポイント.

1. AIを使いこなす上で、特にAIの推論における基盤技術である、自然言語処理の概要を理解する。
2. 1.において、特に複雑な構文構造を持つ自然言語テキストの場合、AIの推論結果に誤りがあり得る事を理解する。
3. AIが推論した内容に関して最終的には人間も結果を確認する事が重要である事を理解する。

症例1、2、4は結果の解釈の際にAIが考慮しない因子やAIが関与しないプロセスについての、症例3、4、5-8、は、AIが活用される「結果の解釈」のプロセスにおける指導ポイントである。これらの指導ポイントは、研究協力者が実際に解釈につまずいた点や、AIを活用したゲノム医療の実装に重要と感じた点である。このことから、少なくとも医学部卒業後数年程度の大学院生や専門家に対しては、これらのポイントを踏まえた、例えば以下の行動目標(案)の様な事項を盛

り込んだカリキュラムの構成にする必要があると考えられた。これらのポイントに共通する事は、AIを活用したゲノム医療を行うには、シーケンス結果の解釈法を学ぶだけでは不十分であり、一連のプロセスの原理を正しく理解する必要があるという事である。

前述した通り、現状ではAIが活用されるのは「結果の解釈」のプロセスのみであり、AIは与えられたシーケンス結果を解析するのみである。結果の判断の主体はあくまでも医師であり、またそれ以外のプロセスの妥当性の評価はできない。AIはあくまでも効率的に情報を提示する支援ツールに過ぎず、そのため例えば、症例2の様にシーケンス手法の選択が適切ではなかった場合、診断や治療方針について誤った結果を提示してしまう危険性がある。またAIが適切なシーケンス結果を解析していても、本事例集では提示しなかったが、解析結果が誤っている場合や、ドラッグ・ラグにより日本では提示された薬剤が使用できない場合も多い。よって専門医はAIの解析結果を鵜呑みにするのではなく、結果を評価し、症例2のように提示された治療を実際に患者に適応可能か、薬剤承認・臨床試験情報や患者背景を考慮して総合的に判断する必要がある。しかしながら、症例3、5で提示した通りAIは短時間で膨大なデータを処理し解析する事

が可能なため、結果の解釈において効率的に有益な情報を得る強力なツールである事は間違いない。AIを活用したゲノム医療を実践の領域にまで推し進める為には、AIによる解釈と人間のそれをどの様に組み合わせるのが適切であるかを事前に検討しておく必要があると考えられる。また症例4、5、6、7で提示した通り、AIを使いこなす上で、特にAIの推論における基盤技術である、機械学習や自然言語処理の仕組み、を事前に理解しておく必要があると考えられる。また、「結果の解釈」ステップの後も、症例6のように、再解析時には、学習によりAIの出力結果が変わり得る事に留意する必要がある。

今回研究協力者となっている医学部卒後数年程度の大学院生は、医学部教育や実臨床の経験から疾患や治療薬等の知識はある程度あるものの、AIやシーケンス、インフォマティクスについての知識は乏しい。また実際に治療に難渋する患者症例を経験しており、ゲノム診療の重要性も実感している。よって一連のプロセスの原理を学び実施し、結果を患者にフィードバックするまで行うカリキュラムは、今回の研究協力者である医学部卒後数年程度の大学院生には効果的であると考えられた。

以上の様に指導事例で挙げた指導ポイントを踏まえ、各患者に合わせゲノム

医療の一連のプロセスを適切に施行・評価することができる知識を持ち、AIの特性を理解した上で有効に活用する事のできる専門医の育成が必要である。これには医学教育だけでは不十分であり、バイオインフォマティクスやAI開発に関わる専門家等の協力が不可欠であるため、多分野と連携した専門医育成の教育基盤整備が急務であると考えられる。

5. 結語

AIを活用したゲノム医療のOJT学習経験を通じて判明した問題点や指導の要点を、具体例を挙げて提示した。この学習経験を通じ、ゲノム医療の一連のプロセスはあくまでも専門家が主体で行う必要があるが、AIは結果の解釈のプロセスにおいて効率的に情報を提示する強力な支援ツールであると考えられた。ゲノム医療の実装にはAIの推論における基盤技術とAIの特性を理解し、有効に活用する事のできる専門医の育成が必要であり、それには多分野と連携した専門医育成の教育基盤整備が急務であると考えられる。

6. 横山班 AIを活用したがんゲノム医療の専門家育成カリキュラム案（概要版）

人工知能(AI)を活用したシーケンス解釈

ねらい：人工知能の判定結果について妥当性を検討でき、人間の解釈を適切に組み合わせる事ができる。

学修目標：

大項目 1. AIが考慮しない因子、プロセスを理解し、説明、実践できる

中項目 1-1. 遺伝医療に関する適切な知識を持ち、研究や検査の適切な説明による適切な同意取得ができる

中項目 1-2. 患者情報(病歴、治療歴、家族歴、患者背景)の適切な収集と解釈ができる

中項目 1-3. 基本的な分子生物学的知識を理解して説明できる。

中項目 1-4. 基本的な分子生物学的操作を理解して実践できる。

中項目 1-5. 従来の塩基検出技術を理解し、主な方法を実践できる

中項目 1-6. 次世代シーケンス法技術を理解できる

中項目 1-7. ライブラリ調整の主な方法を実践できる

中項目 1-8. 塩基配列データとその種類、Informatics原理を理解できる

中項目 1-9. AIが活用される「結果の解釈」のプロセスを理解し、説明できる

小項目 1-9-1. AIの推論における基盤技術の概要を理解できる

小項目 1-9-1-1. 機械学習

小項目 1-9-1-2. 自然言語処理

小項目 1-9-2. 1-8-7-4と関連して、AIが学習する代表的な情報やデータベースについて説明できる

中項目 1-10. AIの特性を人間と対比して理解し説明できる。

中項目 1-11. AIと人間の解釈を適切に組み合わせる事ができる。

中項目 1-12. AIには見逃しがあり得る事を理解できる。

小項目 1-12-1. AIが見逃しがちな情報属性について理解し説明できる。

小項目1-12-1-1.機械学習におけるパスウェイターゲットでないaction

小項目1-12-1-2.自然言語処理に起きる複雑な構文構造をもつテキスト

中項目 1-13. 学習により、AIの出力内容や推論精度が変わり得る事を理解する。

中項目 1-14. AIの推論内容に関して人間も結果を確認する必要性を理解し、

実践できる。

中項目 1-15. 1-1.と関連して、正常対照の解析に伴い、偶発的所見として生殖細胞系列変異が認められる場合がある事をその対応と共に理解し実践できる。

7. 横山班 AIを活用したがんゲノム医療の専門家育成カリキュラム案（詳細版）

人工知能(AI)を活用したシーケンス解釈

ねらい：人工知能の判定結果について妥当性を検討でき、人間の解釈を適切に組み合わせる事ができる。

学修目標：

大項目 1. AIが考慮しない因子、プロセスを理解し、説明、実践できる

中項目 1-1. 遺伝医療に関する適切な知識を持ち、研究や検査の適切な説明による適切な同意取得ができる

小項目 1-1-1. ヒトゲノム、遺伝子の構造と機能

小項目 1-1-1-1. 5'非翻訳領域(5'UTR)

小項目 1-1-1-2. 3'非翻訳領域(3'UTR)

小項目 1-1-1-3. センス鎖とアンチセンス鎖

小項目 1-1-1-4. 遺伝子名

小項目 1-1-1-4-1. Human Genome Organisation, HUGO と HUGO Gene Nomenclature Committee (HGNC)

小項目 1-1-1-5. 遺伝子の発現

小項目 1-1-1-5-1.転写

小項目 1-1-1-5-1-1. 非転写鎖/転写鎖

小項目 1-1-1-5-2.プロモーター

小項目 1-1-1-5-3.調節因子

小項目 1-1-1-5-4.転写因子

小項目 1-1-1-5-5.RNAプロセッシングとスプライシング

小項目 1-1-1-5-6.選択的スプライシング

小項目 1-1-1-5-6-1.Canonical transcriptについて

小項目 1-1-1-5-7.mRNA安定性

小項目 1-1-1-5-8.翻訳

小項目 1-1-1-5-8-1.遺伝コード、コドン表、蛋白の略号

小項目 1-1-1-5-8-2.タンパク質の翻訳後プロセッシング

小項目 1-1-1-5-8-3.タンパク質の分解

小項目 1-1-1-5-9.遺伝子発現調節に関わるエピジェネティック修飾変化

小項目 1-1-1-5-9-1.DNA塩基のメチル化による遺伝子発現の変化

小項目 1-1-1-5-9-1-1CpGアイランド

小項目 1-1-1-5-10.ヒストンの化学修飾による遺伝子発現の変化

小項目 1-1-1-5-10-1.ヒストン修飾とその機能：ヒストン暗号仮説

小項目 1-1-1-5-10-2.ヌクレオソームとヒストンタンパク質(H2A、H2B、
H3、H4)

小項目 1-1-1-5-10-3.ヒストンアセチル化酵素 (histone acetyl transferase:
HAT)

小項目 1-1-1-5-10-4.ヒストン脱アセチル化酵素 (histone deacetylase:
HDAC)

小項目 1-1-1-5-10-5.ヒストンメチル化酵素 (histone methyltransferase:
HMT)

小項目 1-1-2. 遺伝子異常と疾患

小項目 1-1-2-1.単一遺伝子疾患と多因子疾患の違い

小項目 1-1-2-2.代表的な遺伝性疾患の分類と頻度

小項目 1-1-2-3. 単一遺伝子疾患

小項目 1-1-2-3-1.メンデル遺伝学

小項目 1-1-2-3-2.遺伝型

小項目 1-1-2-3-3.表現型

小項目 1-1-2-3-4.ホモ接合体

小項目 1-1-2-3-5.ヘテロ接合体

小項目 1-1-2-3-6.優性と劣性

小項目 1-1-2-3-7.保因者

小項目 1-1-2-3-8.常染色体優性遺伝病

小項目 1-1-2-3-8-1.浸透率

小項目 1-1-2-3-8-2.不完全浸透

小項目 1-1-2-3-8-3.新生突然変異

小項目 1-1-2-3-9.常染色体劣性遺伝病

小項目 1-1-2-3-10.X連鎖劣性遺伝病

小項目 1-1-2-3-11.新生突然変異

小項目 1-1-3. 集団遺伝学

小項目 1-1-3-1.遺伝の変異

小項目 1-1-3-2.単一塩基置換

小項目 1-1-3-3.欠失

小項目 1-1-3-4.挿入

小項目 1-1-3-5.遺伝子重複

小項目 1-1-3-6.遺伝子融合

小項目 1-1-3-7.遺伝的多型

小項目 1-1-3-7-1.マイナーアレルとメジャーアレル

小項目 1-1-3-7-2.SNP(一塩基多型)

小項目 1-1-3-7-3.コピー数多型(CNV)

小項目 1-1-3-7-4.ミニサテライト多型・マイクロサテライト多型 (STR)

小項目 1-1-3-7-5.ハーディワインベルグの法則

小項目 1-1-3-7-6.生物の進化

小項目 1-1-3-7-6-1.突然変異

小項目 1-1-3-7-6-2.分子時計と近縁性推定

小項目 1-1-3-7-6-3.遺伝的距離

小項目 1-1-3-7-7.薬剤反応性

小項目 1-1-3-8.薬剤反応性

小項目 1-1-3-9.変異や多型の記載法

小項目 1-1-3-9-1. Human Genome Variation Society, HGVS

小項目 1-1-3-9-2. その他

小項目 1-1-2. 遺伝子検査とその目的、診断

小項目 1-1-3. 家系図作成と遺伝カウンセリング

小項目 1-1-4. チーム医療

小項目 1-1-5. 遺伝子治療

小項目 1-1-6. 倫理

小項目 1-1-6-1. 遺伝倫理

小項目 1-1-6-2. 遺伝情報管理、個人情報の匿名化

**中項目 1-2. 患者情報(病歴、治療歴、家族歴、患者背景)の適切な収集と解釈
ができる**

小項目 1-2-1. 病歴

小項目 1-2-2. 治療歴

小項目 1-2-3. 家族歴と家系図作成

小項目 1-2-4. 染色体検査など核型分析結果

小項目 1-2-4-1. 染色体検査の基礎知識と分染法の原理

小項目 1-2-4-2. 染色体の基本的構造

小項目 1-2-4-3. 核型分析とその記載法(国際基準ISCN2016)

小項目 1-2-4-3-1. 構造異常

小項目 1-2-4-3-1-1. 転座

小項目 1-2-4-3-1-1. 欠失

小項目 1-2-4-3-1-1. 挿入

小項目 1-2-4-3-1-1. 重複

小項目 1-2-4-3-1-1. 逆位

小項目 1-2-4-3-2. 数的異常

小項目 1-2-4-3-2-1. トリソミー

小項目 1-2-4-3-2-1. モノソミー

小項目 1-2-4-3-2-1. ナリソミー

小項目 1-2-4-3-3. クローン性の判定

小項目 1-2-5. その他の検査法

中項目 1-3. 基本的な分子生物学的知識を理解して説明できる.

中項目 1-4. 基本的な分子生物学的操作を理解して実践できる.

小項目 1-4-1. マイクロピペットの扱い

小項目 1-4-2. 検体前処理 (溶血、細胞分離、凍結破碎、脱パラフィン処理、ホモジナイザーを用いたホモジナイズ)

小項目 1-4-3. 核酸抽出(DNA/RNA)

小項目 1-4-4. 核酸濃度測定 (分光光度計、蛍光濃度測定)

小項目 1-4-5. 微量サンプルの核酸の純度、濃度測定

小項目 1-4-5-1. Agilent社 2100 バイオアナライザ

小項目 1-4-5-2. Agilent社 TapeStation

小項目 1-4-6. 核酸精製 (カラム、エタノール置換、磁気ビーズ法)

小項目 1-4-7. PCR法などの遺伝子増幅法

小項目 1-4-7-1. 意義と原理

小項目 1-4-7-2. プライマー、プローブ合成及びPCR反応条件

小項目 1-4-7-3. PCR法

小項目 1-4-7-4. アガロースゲル電気泳動

小項目 1-4-7-5. RT-PCR法

小項目 1-4-7-6. リアルタイムPCR法

小項目 1-4-7-7. その他の遺伝子増幅技術(LAMP、TMA、デジタルPCR法等)

中項目 1-5. 従来の塩基検出技術を理解し、主な方法を実践できる

小項目 1-5-1. サンガーシーケンス法

小項目 1-5-2. その他の従来の検出法

中項目 1-6. 次世代シーケンス法技術を理解できる

小項目 1-6-1. 意義と塩基配列決定の原理

小項目 1-6-2. ショートリードシーケンサー

小項目 1-6-2-1. 可逆的ターミネーター色素標識ヌクレオチドを用いたSBS (Sequencing-by-Synthesis) 法

小項目 1-6-2-2. 塩基取り込み時に水素イオンの放出を測定する手法

小項目 1-6-3. ロングリードシーケンサーなどその他の手法

中項目 1-7. ライブラリ調整の主な方法を実践できる

小項目 1-7-1. インプットDNAの定量と品質評価

小項目 1-7-2. DNAの断片化

小項目 1-7-3. DNAの末端修復と3'末端のアデニル化

小項目 1-7-4. インデックス(分子バーコード)アダプターとDNA断片のライゲ

ート

小項目 1-7-5.アダプター付きDNA断片の増幅

小項目 1-7-6.アダプター付きDNA断片の精製

小項目 1-7-7.アダプター付きDNA断片の定量と品質チェック

小項目 1-7-8.アダプター付きDNA断片のハイブリダイゼーション

小項目 1-7-9.ハイブリダイゼーションDNAのキャプチャーと磁気ビーズによるサイズ選択的精製

小項目 1-7-10.キャプチャー後ライブラリの増幅

小項目 1-7-11.ライブラリ精製

小項目 1-7-12.ライブラリの定量と品質チェック

小項目 1-7-13. シークエンス

中項目 1-8. 塩基配列データとその種類、Informatics原理を理解できる

小項目1-8-1. シークエンスデータ

小項目1-8-1-1.fastq (テキストデータ)

小項目1-8-1-2.fast.gz (圧縮)

小項目1-8-2. アラインメントデータ

小項目1-8-2-1.BAM (容量を減らす為にバイナリ変換された物)

小項目1-8-3.変異データ

小項目1-8-3-1.VCF

小項目1-8-4.その他の形式のデータ

小項目1-8-5.塩基配列決定精度

小項目1-8-6. 解析に必要なデータ量

小項目1-8-7. ヒト参照ゲノム配列

小項目1-8-7. 主な解析パイプラインと解析フロー

小項目1-8-7-1. アダプター情報の除去とクオリティコントロール

小項目1-8-7-2.配列アラインメント, マッピング

小項目1-8-7-3. Fischer検定による統計的仮説検定(変異コール)

小項目1-8-7-4. 遺伝子アノテーションや代表的なデータベース

小項目1-8-7-4-1. ゲノムデータベース

小項目1-8-7-4-1-1. UCSC Genome Browser

小項目1-8-7-4-1-2. その他

小項目1-8-7-4-2. SNP(一塩基多型)データベース

小項目1-8-7-4-2-1. Human Genetic Variation Database (HGVD)

小項目1-8-7-4-2-2. ToMMo

小項目1-8-7-4-2-3. dbSNP

小項目1-8-7-4-2-4. the 1000 Genomes Project

小項目1-8-7-4-2-5. gnomAD

小項目1-8-7-4-2-6. ExAC

小項目1-8-7-4-2-7. その他

小項目1-8-7-4-3. 癌の体細胞変異のデータベース

小項目1-8-7-4-3-1. COSMIC

小項目1-8-7-4-3-2. cBioportal

小項目1-8-7-4-3-3. ICGC

小項目1-8-7-4-4. 遺伝性疾患のデータベース

小項目1-8-7-4-4-1. OMIM

小項目1-8-7-4-4-2. その他

小項目1-8-7-4-5. 薬剤、添付文書データベース

小項目1-8-7-4-6. パスウェイデータベース

小項目1-8-7-4-6-1. KEGG

小項目1-8-7-4-7. タンパク質のドメインデータベース

小項目1-8-7-4-8. タンパク質の機能予測アルゴリズム

小項目1-8-7-4-9. 治験情報データベース

小項目1-8-7-4-9-1. 臨床研究実施計画・研究概要公開システム (jRCT)

小項目1-8-7-4-9-2. 大学病院医療情報ネットワーク研究センター (UMIN-CTR)

小項目1-8-7-4-9-3. JAPIC (Japic Clinical Trials Information/JapicCTI)

小項目1-8-7-4-9-4. clinicaltrials.gov

小項目1-8-7-4-10. FDA承認薬に対応する遺伝子変異データベースやナレッジ

ベース

小項目1-8-7-4-10-1. OncoKBとエビデンスレベル (Levels of Evidence)

小項目 1-8-7-4-10-2. PharmGKB

小項目1-8-7-4-10-3.CIViC

小項目1-8-7-4-11. その他のツールやデータベース、GWAS等

中項目 1-9. AIが活用される「結果の解釈」のプロセスを理解し、説明できる

小項目 1-9-1. AIの推論における基盤技術の概要を理解できる

小項目 1-9-1-1. 機械学習

小項目 1-9-1-2. 自然言語処理

小項目 1-7-1-3. その他のAIの基盤技術

小項目 1-9-2. 1-8-7-4と関連して、AIが学習する代表的な情報やデータベースについて説明できる

中項目 1-10. AIの特性を人間と対比して理解し説明できる。

中項目 1-11. AIと人間の解釈を適切に組み合わせる事ができる。

中項目 1-12. AIには見逃しがあり得る事を理解できる。

小項目 1-12-1. AIが見逃しがちな情報属性について理解し説明できる。

小項目1-12-1-1.機械学習におけるパスウェイターゲットでないaction

小項目1-12-1-2.自然言語処理に起きる複雑な構文構造をもつテキスト

小項目1-12-1-3. その他

中項目 1-13. 学習により、AIの出力内容や推論精度が変わり得る事を理解する。

**中項目 1-14. AIの推論内容に関して人間も結果を確認する必要性を理解し、
実践できる。**

中項目 1-15. 1-1.と関連して、正常対照の解析に伴い、偶発的所見として生殖細胞系列変異が認められる場合がある事をその対応と共に理解し実践できる。

7. 健康危険情報

特記事項なし

8. 研究発表

論文発表

1. Hijikata K, Yokoyama K (co-1st and co-corresponding author), Yokoyama N, Matsubara Y, Shimizu E, Nakashima M, Yamagishi M, Ota Y, Lim L, Yamaguchi R, Ito M, Tanaka Y, Denda T, Tani K, Yotsuyanagi H, Imoto S, Miyano S, Uchimar K, Tojo A. Successful clinical sequencing by molecular tumor board in an elderly patient with refractory Sézary syndrome. JCO Precis Oncol 2020 in press.
2. 横山和明. 第 81 回日本血液学会学術集会 Presidential シンポジウム論文 血液がんのプレシジョン・メディシンにおける AI 活用の試み 臨床血液 2020 61(5):1-10. in press
3. Takeda R, Yokoyama K (corresponding author), Kobayashi S, Kawamata T, S Nakamura, Tomofusa F, Ito M, Yusa N, Shimizu E, Ohno N, Yamaguchi R, Imoto S, Miyano S, Uchimar K, and Tojo A. An Unusually Short Latent Period of Therapy-Related Myeloid Neoplasm Harboring a Rare MLL-EP300 Rearrangement: Case Report and Literature Review Case Rep Hematol, 2019, 4532434 doi.org/10.1155/2019/4532434
4. Nakamura S, Yokoyama K (co-1st and co-corresponding author), Shimizu E, Yusa N, Kondoh K, Ogawa M, Takei T, Kobayashi A, Ito M, Isobe M, Konuma T, Kato S, Kasajima R, Wada Y, Inoue-Nagamura T, Yamaguchi R, Takahashi S, Imoto S, Miyano S, and Tojo A. Prognostic impact of circulating tumor DNA status post-allogeneic hematopoietic stem cell transplantation in AML and MDS. Blood. 2019; 133(25):2682-2695

9. 学会発表

一般口頭発表

1. 横山和明, 血液疾患のプレシジョンメディシンにおける AI 活用の試み, 日本血液学会総会・Presidential Symposium, 東京国際 Forum, 2019/10/11
2. 横山和明, 腫瘍由来循環 DNA を用いた AML および MDS の移植後微小残存病変検出に関する後方視的解析, 神奈川移植 Forum, 中外製薬 神奈川支店, 2019/11/30,

3. 近藤 幹也, 横山 和明, 遊佐 希, 伊藤 美香, 清水 英悟, 中村 聡介, 小川 弥穂, 武井 智美, 小林 麻子, 笠島 理加, 和田 結花, 山口 類, 井元 清哉, 長村 登紀子, 内丸 薫, 宮野 悟, 東條 有伸 治療関連骨髄系腫瘍 15 例のゲノム解析と液体生検によるファウンダークロンの後方視的追跡 第 81 回日本血液学会学術集会 東京国際 Forum 2019/10/13
4. 小川 弥穂, 横山 和明, 塚田 信弘, 余語 孝夫 3, 近藤 幹也, 武井 智美, 中村 聡介, 伊藤 美香, 小林 麻子, 遊佐 希, 山本 菜子, 笠島 理加, 清水 英悟, 山口 類, 井元 清哉, 宮野 悟, 東條 有伸 臨床シーケンスにて確定診断に至った慢性好中球性白血病の 2 例 第 81 回日本血液学会学術集会 東京国際 Forum 2019/10/13

ポスター発表

1. Kazuaki Yokoyama, Nozomi Yusa, Mika Ito, Miho Ogawa, Kanya Kondoh, Eigo Shimizu, Seiya Imoto, Arinobu Tojo, Satoshi Takahashi "Cell free-DNA based detection of minimal residual disease post allogenic-stem cell transplantation in acute myeloid leukemia and myelodysplastic syndromes" 7th US-Japan Workshop on Biomarkers for Cancer Early Detection (AMED-NCI) 東京大学 伊藤謝恩ホール 2020/1/27
2. 近藤幹也 治療関連骨髄系腫瘍 15 例のゲノム解析と液体生検によるファウンダークロンの後方視的追跡 第 2 回 COI 学会 JST 2019/9/20
3. "Tomomi Takei, Kazuaki Yokoyama, Sousuke Nakamura, Miho Ogawa, Kanya Kondoh, Eigo Shimizu, Rika Kasajima, Rui Yamaguchi, Seiya Imoto, Satoru Miyano, Arinobu Tojo" "The utility of exome sequence of circulating tumor DNA in drug-resistant and/or advanced phase chronic myeloid leukemia" The 78th Annual Meeting of the Japanese Cancer Association 京都国際会館 2019/9/28

その他の発表

特記事項なし

知的所有権の取得状況

1. 特許取得
特記事項なし
2. 実用新案登録
特記事項なし
3. その他
特記事項なし