

新春
放談

検査・診断領域における AIの現状と未来



2050

2040



2030



2026

2020



2000



1990

語り手

石川 俊平 先生

東京大学大学院医学系研究科 衛生学分野 教授

奥 真也 先生

東京科学大学 医療・創薬イノベーション教育開発機構 特任教授

下澤 達雄 先生

国際医療福祉大学大学院医学研究科 臨床検査医学 主任教授

柳原 克紀 先生

長崎大学大学院医歯薬学総合研究科 病態解析・診断学分野 教授

(五十音順)

聞き手

堤 武也 先生

東京大学大学院医学系研究科 病因・病理学専攻感染制御学分野
内科学専攻生体防御感染症学分野 教授/モダンメディア編集委員

大西 宏明 先生

杏林大学医学部 臨床検査医学教室 教授/モダンメディア編集委員

はじめに

堤 本日、進行を務めさせていただきます東京大学の堤と申します。一緒に杏林大学の大西先生に司会をしていただきます。前半は私が進行させていただきます。後半は大西先生にお願いする予定です。今回、検査領域におけるAIということでテーマを決めて座談会を開くことになりました。今回この座談会を開くにあたり編集委員のほうでテーマを検討しましたが、かなりAIが進歩してきていますので、この辺りで一度AIをとりあげたほうが良いのではないかとということで、ほぼ満場一致で承認され、開催させていただくことになりました。

このような形で、検査領域の中で、それぞれ専門領域で非常にご活躍、ご高名な先生にお話をいただくことを非常に光栄と思っております。

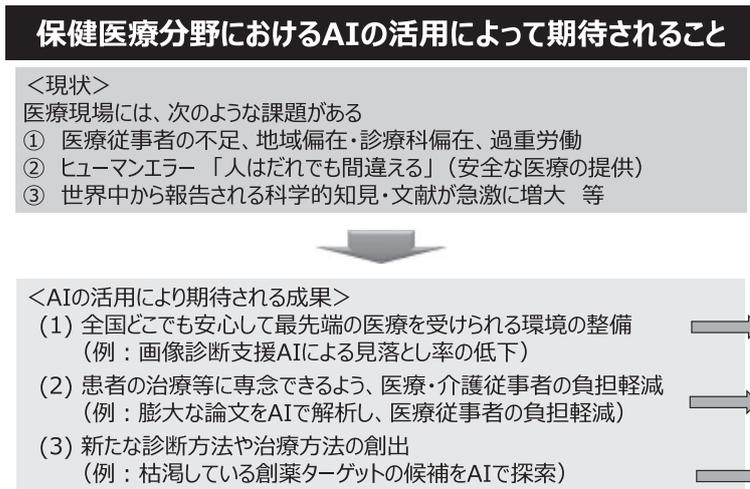
私からの簡単なイントロで、釈迦に説法かもしれませんが、2018年の厚労省の資料から引用してまいりました(図1)¹⁾。医療現場における課題ということで、医療従事者の不足、地域偏在・診療科偏在、過重労働といった問題。それからヒューマンエラー、それから科学的知見や文献が急激に増大しているという課題を基にAIを活用することが厚労省で議論されています。



堤 武也 先生

AIの活用により期待される成果として、(1) 全国どこでも安心して最先端の医療を受けられる環境の整備。これは診断精度の向上、医療過誤の防止、地域格差の是正につながるかと思います。(2) 患者の治療等に専念できるよう、医療・介護従事者の負担軽減ということで、医師の負担軽減、業務の効率化につながるものかと思われます。また、(3) 新たな診断方法や治療方法の創出ということで、医療の進歩につながるようなことが提案されています。

その際に、同じ資料からですが、AI開発を進めるべき重点6領域ということがこのような形で提案されています(図2)¹⁾。詳細は省きますが、ゲノ



出典：厚生労働省、資料2「保健医療分野におけるAI開発の方向性について」(2018年7月23日)
<https://www.mhlw.go.jp/content/10601000/000337597.pdf>

文献1) より転載、一部改変

図 1

ム医療、画像診断支援、診断・治療支援、医療品開発、医薬品開発、介護・認知症、手術支援ということになります。本日は画像診断支援ならびに診断治療支援のところを活発に議論させていただければと思います。

このような背景の下、さまざまなところで診断領域のAIの開発がなされて、私が調べた限りだと、現在保険承認されているものとして、インフルエンザ診断の nodoca、こちらについては後ほど柳原先生からお話をいただきます。それから、本日は内視鏡関連の先生をお呼びしていませんが、大腸内視鏡診断支援ということで、EndoBRAIN-EYEをはじめとして保険算定できるものがあるかと思います。EndoBRAIN-EYEは、大腸内視鏡の際にポリープをAIが検知してくれるようなものようです。

医療におけるAI活用のメリットとして、先ほどお話したようなものがあります。一方、問題点として、精度や信頼性の問題、AIが誤診した場合に誰が責任を取るかといった責任の所在、データプライバシー・セキュリティ、過度な依存によって医療従事者自体のスキルが低下するといった懸念もあります。こういったところについて先生方からお話やご意見をいただければと思います。

私のイントロは以上です。早速ですが、東京科学大学の奥真也先生からお話をいただきたいと思い



奥 真也 先生

ます。奥先生からは、AIと医療について総論的なお話をいただき、可能であれば先生のご専門領域でもある放射線診断領域についてもお話をいただければと思います。奥先生、よろしく願いいたします。

AIと医療 歴史・課題・行きつく先

奥 よろしく願いいたします。私は放射線科医で、東大に2003年までいて、その後に埼玉医大に2009年までいました。

その間に東大の22世紀医療センターで健診情報

懇談会では、次の両面からAI開発を進めるべき重点6領域を選定。

- ①我が国における医療技術の強みの発揮
 - ②我が国の保健医療分野の課題の解決（医療情報の増大、医師の偏在等）
- これら6領域を中心に、AIの研究開発を加速化させる。

【AIの実用化が比較的早いと考えられる領域】

領域	我が国の強み（○） /課題（△）	AIの開発に向けた厚生労働省の主な施策 （民間企業におけるAI開発を促進するための基盤を整備）
①ゲノム医療	△欧米に比べて取組が遅れ	・ 国立がん研究センターにがんゲノム情報管理センターを整備し、ゲノム情報を集約 ・ がんゲノム情報管理センターが臨床情報や遺伝子解析情報等を横串で解析する知識データベースを構築
②画像診断支援	○日本の高い開発能力 ○診断系医療機器の貿易収支も黒字（1,000億円）	・ 関連医学会（日本病理学会、日本消化器内視鏡学会、日本医学放射線学会、日本眼科学会）が連携して画像データベースを構築 ・ 厚生労働省が、医師法上や医薬品医療機器法上の取組を明確化
③診断・治療支援 （問診や一般的検査等）	△医療情報の増大によって医療従事者の負担が増加 △医師の地域偏在や診療科偏在への対応が必要 △難病では診断確定までに長い期間	・ 日本医療研究開発機構（AMED）研究費により、難病領域を幅広くカバーする情報基盤を構築 ・ 厚生労働省が、医師法上や医薬品医療機器法上の取組を明確化
④医薬品開発	○日本は医薬品創出能力を持つ数少ない国の1つ ○技術貿易収支でも大幅な黒字（3,000億円）	・ 国立研究開発法人 医薬基盤・健康・栄養研究所が、創薬ターゲットの探索に向けた知識データベースを構築 ・ 国立研究開発法人 医薬基盤・健康・栄養研究所、理化学研究所、及び京都大学が中心となり、製薬企業とIT企業のマッチングを支援

【AIの実用化に向けて段階的に取り組むべきと考えられる領域】

⑤介護・認知症	△高齢者の自立支援の促進 △介護者の業務負担軽減	・ 厚生労働科学研究費補助金により、介護における早期発見・重症化予防に向けたデータ収集及び予測ツールの開発
⑥手術支援	○手術データの統合の取組で日本が先行 △外科医は数が少なく、負担軽減が必要	・ 厚生労働科学研究費補助金等により、手術関連データを相互に連結するためのインターフェースの標準化を実施

出典：厚生労働省、資料2「保健医療分野におけるAI開発の方向性について」（2018年7月23日）
<https://www.mhlw.go.jp/content/10601000/000337597.pdf>

文献1) より転載

の標準化、健診情報学講座という NTT データがつくった寄附講座がありまして、そこの特任准教授を5年間やって、その後、会津大学という工学系の福島にある大学に移り、直近十何年は医療機器の会社や製薬企業、外資ばかり4社を経験して、今年の3月に東京科学大学特任教授でアカデミアに戻りました。

それも今日の話とは関係ないですが、自分で研究費を集めて、そのお金でポストを作ってもらって就職したということなので、既存の組織とはあまり関係のない形で医療介護 DX の研究をしています。幸いに初年度から AMED のお金が取れたので、もう少し暇な予定だったのですが、急に忙しくなりました。

今日は本を持ってきたのでご参考にですが、放射線科医としてよりも、企業にいる医師として働き始めてから病院で見ている風景とは全然違うものが見えてきたので、そういうような経験をいろいろもとにし、あとは薬事のコンサル会社にもいましたので、そういう経験ももとに、『未来の医療年表』という本を講談社から出しました。これが5年前。本の表紙に書かれている「初の本格的認知症薬誕生」と、これは2024年にレカネマブが出たので、当たっているといえは当たっている。もちろん、先生方もご存じのように治験というものは公開して進行させているので、それは予想というよりは既定の事実を書

いているのに近いですが、それでも2020年に本を書いて、今年2025年の状況は予測できていました。そのときに2030年はAI診療が主流にということで予想していました。なぜかという話はこの後にさせていただきます。

また、今年9月に『AIに看取られる日』を出したばかりですが、いろいろな医療のプロセス全てにAIが関係していき、それだけではなくて、人間医師の独壇場にみえる診察室もそうですし、診断の分野には大きく入り込んできている。最後は死を看取るところまでいきそうということで、その辺の、そう、啓発活動というとおこがましいですが、そういうこともやらせていただいています。

象徴的なスライドなので持ってきたのですが(図3)、私がこうなるだろうと思って最近の著書に書きました。これは4者モデルと私は呼んでいます。お医者さんと患者さんの間にAIが2つあって、患者さん側のAIと医師側のAIと両方あって、それぞれが代弁する。下のほうがヒツジとワンちゃんの形をしています。この2者が一番濃厚な話をする。

よくある話で、お医者さん側は3分でやりたい、患者さん側は3時間話を聞いてほしいと。その60倍の時間の長さギャップをどうやって埋めるかということで、医学の知識や単語の使い方もそうですし、あとはスピードですよね。皆さんも高齢のお父さんやお母さんとお話するときゆっくり話さなくて



図3

はいけなくて、自分は時間がないのということをご経験があるかと思いますが、そういうところも含めて埋めることが今後必要になってくる。この4者が一堂に会する必要があるかどうかはまた別の議論なので、これはあくまでも象徴的なスライドですが、こうやって4者がそれぞれ会話するような形に今後必ずなります。断言してもいいような状況で、1、2年でどこまで行くか。2024年から25年にAIは非常に進んでいるので、これは現実的な話だと受け止めていただけて結構です。

制度の問題、仕組みのほうの問題として、それがなかなか実現できないことはいろいろ出てくる。先ほど問題提起された責任問題。誰が責任を持つかという問題も当然あると思います。

AIのヒストリーもお話に入れていただくよう編集室事務局からご指示を受けましたので、資料に入れました(図4)。極々分かりやすくまとめましたが、1956年から90年の、先生方がお若い頃にお聞きになった頃から今まで。先ほど座談会が始まる前に森兼啓太先生の名前が出ましたが、森兼先生と一緒に東大の赤レンガの中に研究室があって、当時開原先生や大橋先生、大江和彦先生に指導をいただきながら、学生としてコンピュータの研究をしていました。

その頃のちょっと前、1956年だから、われわれが生まれるちょっと前ぐらいに、概念としてはAIは提唱されました。ただ、そのときは記号処理というか、こういうルールで処理することを人間が決め

て、それをコンピュータが演繹するというだけだったのでなかなかうまくいかなかった。

それで Recursive Neural Network というコンピュータ的なシステムを作って、そこで次に出てくる単語は何かという話を再帰性、Recursive Neural Network というもので行われたものなのです。図の右側がそれを象徴したものですが、これはおとぎ話です。Once upon a time, there lived an old man とくると、次には and an old woman がくるだろうと。おじいちゃんといったら、おばあちゃんに決まっているでしょう。そういう話を持ってきてだけです。

今やっているAIも原理をたどれば同じことをやっているだけ。次に来る単語はこれが一番もってもらいたいというのを。昔はルールでやっていただけ、今はそれをものすごく大量のデータでやるということになっている状況です。2行目に書きましたように、コンピュータCPUの力が強くなって、ビッグデータでデータがたくさんある。その二つの組み合わせで勝ち筋というか、そこでちゃんといけるんじゃないかという話が出てきて、AI = 深化学習の流れになる。

私は囲碁が趣味なのですが、2017年に Deep Mind という Google の傘下にいる会社が作ったAIが、韓国のイ・チャンホという囲碁の世界チャンピオンに5回やって4回勝ってしまったのです。まるで勝負にならなかった。囲碁は皆さんおやりになるか分かりませんが、8年前は互戦で、ハンディキャップなしで4勝1敗ということは、これは勝てないなど

AI Chronography

1956-1990

1956に概念提唱。記号処理AIの限界と、RNN・ベイズ統計・SVMなど統計的AIによる局所最適の限界が続く

2012

GPU+BigDataによる勝ち筋が見え、AI=進化学習の流れに

2017

Deep Mindの囲碁世界チャンピオン撃墜による社会的認知

2024

静かに到達した初期 Singularity

Recursive Neural Network

Once upon a time,
there lived an old
man and an old
woman, who....

図4

いう感じ。その前をたどると将棋が負けるようになって、そのもっと前にチェスが負けるようになって、人類がどんどん克服されていっている。

私どもが学生だった頃には、数学の話ですがNP有界問題という、数学的に場合の数が多すぎるという問題によって、この問題は21世紀中には解決しないとされていました。22世紀には変化が起きてくるけれど、21世紀の間は人間はまだ大丈夫、コンピューターには追いつかれない、とずっと言われていた。

だけど全然早くなってしまって、今や人間は負ける状態。Deep Mindが囲碁の世界で、2017年に並んだのですけれど、2025年の今は三子局でも勝てない。最初に石を3個置いて人間は全然勝てない、ということです。プロのトップがそんな状態になっている。そのぐらい8年間で差がついてしまいました。

私が勝手に言っているのですが、2024年には静かにSingularityの初期に到達した。Singularityはこの分野ではよく使われる言葉で、コンピューターの性能が高くなってきて、ビッグデータの恩恵もあって、人間を超える知的上層活動を能動的にやるようになることを普通は呼んでいますけれども、それが、2024年後半ぐらいから見えてきているという怖い時代な感じですよ。

Kurzweilという人はSingularityでは有名な学者で、この人が2024年のインタビューで、今だよということを言っています。それから象徴的な論文を1本だけご紹介すると、arXivなので査読をまだ受

けていない論文ですが、ご存じのようにコンピュータの世界は査読が間に合わないの、査読が間に合わないで出てくるのがいっぱいあります。

2024年、Jinさんという中国人の人が書いた論文で、去年にはthe fastest point of the current AI waveということで、一番急峻にAIが進歩するのは2024年だと論文にも書いてあって、私もそれに一票を投じているほうです。

2024年から25年に既にSingularityに来ていると私は考えています。多くの論文を見ていただくと、2035年から40年ぐらいに来ると言っている人が多いです。多数決を取ると。でも、私は既にその段階に入っていると個人的には考えています。

AIにできることをまとめており(図5)、左側は国策で、国がAIやDXで何をできるかということで、厚労省やデジタル庁、内閣官房が何を目標にしているかという図になります。診療情報の標準化、処方箋、医療介護データを一括管理ということをしていて、これは結構今進んでいます。けど、そんなものではないです。それは国が遅れているというか、誰もついていけないので、そんな状況になっています。

図の右側にありますように、まずは人間がやっている医療の代替はできるようになっていて、診断代替、ガイドラインに基づいて治療のオプションを提供するところは完全にできています。先ほど放射線科のことを話すように言われましたが、放射線科の領域はもう勝てないです。放射線科医はその意味に

AIができること

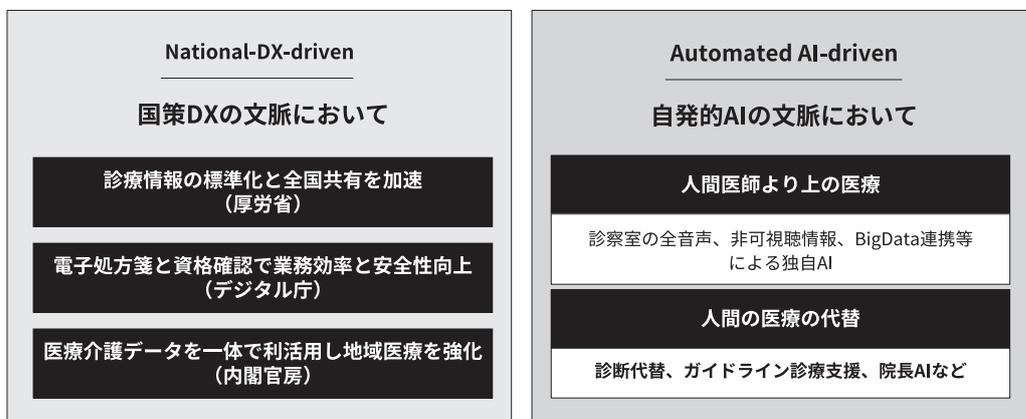


図5

おいては「終わり」と思っています。放射線科と病理は技術的には終わりなのですけど、ただ制度的にお医者さんが最後は判子を押さなければいけないなどいろいろあって、なくならない。東大放射線科教授の阿部 修先生にも「奥先生、あまり放射線科医は終わりなんて言わないように」といつも言われているのですけれど、もう技術的には、終わりです。

図の右側、人間医師より上の医療を見ていただくと、これは診察室の全音声を取って、その音声を基に電子カルテのカルテ情報に反映するのは、たぶん今年ぐらいにプロトタイプがいくつか出てきている。例えば介護。おじいちゃん、おばあちゃんが職員側において、老老介護状態。職員もお年寄り、介護される人もお年寄りなので、例えば、記録の入力が難しい。そこで、需要があるのでわりと早く進んでいます、音声でそのときの様子を全部取っている。プライバシーの問題はありますが、取っておいたらカルテができてしまう。そこから情報を得る。

もう一つ、まだ資料として、いろいろと提示するものがないのですが、非可視聴と書きましたが、例えば人間がX線診断をするときはフィルムを読むわけですから視覚情報です。そうすると紫外線や赤外線、要するに人間の可視光を見る目では見えない情報は見えていないです。音も一緒に、診察で、耳で聞こえる音は聞こえるのですが、耳で聞こえないような低音やモスキート音のようなものは診断に使われていない。匂いも使われてない。そういうものの中に診断に組みする情報は絶対あるはずなんです。なので、そこは人間より上の医療がこれから出てくるのは確実だと思います。

そして、もう一つだけ、内視鏡の話などいろいろ出てきましたが、日本は診断アプリ、治療アプリはむちゃくちゃ遅れています。先ほどご説明いただいたものや肥満治療、禁煙治療など、五指に余らないぐらいのものしかソリューションが出ていない。海外はものすごく進んでいて、アメリカやカナダ、ヨーロッパもものすごく進んでいる。使われている数が全く違う。そこは個人情報保護法やさまざまな日本の法律によって止まっている部分があると理解いただければいいと思います。

大学で研究していると言いましたが、大学でどういう研究を私がしているかということで、ちらっと

ですが、いろいろな医療機関、東大病院もそうですし、クリニックもそうですが、大きな期待がDXやAIにはあります。けれども、その中で本当にやることはどのぐらいのものなのかという低めのストライクゾーンで目標を決めるのが大学の研究、最小セットというのを研究しようとしています。

いろいろなところに手伝ってもらって、ここはいちいち触れませんが、そこで医療DXは、本当はこうあるべきだと。DXとAIはちょっと違うのですが、DXがAIを含むと考えていただいとりあえずは結構だと思います。

みんながどのようなものを望んでいるかについて、クリニックの先生方にだいぶインタビューをしましたが、皆さんAI-drivenな新しいテクノロジーを早く入れたくてしょうがない。AIで問診を代わりにやってくれるもの。事前問診をやるのは、例えば東大の後輩で、われわれより下の世代のUbieという会社があって、社長の阿部 吉倫先生のところがAI問診をかなりやっています。

そういうものや、診療Chatbotで、先生のところに来る前に初期的な、これ風邪だと思うけれど行ったほうがいいですかねみたいな話はChatbotで済ませてしまう。あとは、クリニックのそれぞれの院長先生が、この先生は何を言いそうというのを、その人の講演録や過去の患者さんデータから、きっとこの人だとこの薬はこうだよとか、ここは薬を飲まなくてもいいんじゃないのとか、お医者さんがどういことを言うかということも結構ソリューションが出始めています。

そういう上記で話をした点と、最近言われているような人件費、人が足りない問題に対してやってほしいと。これはどちらかというとテクノロジーとしては下のほうになります。AI-drivenのものと人件費節約のものが皆さん欲しい。今は、人がそもそも取れない。

あとは死への踏み込み。AIが死に関わる医療まで代替していいかというレベルのコンセンサスは、私の読みでは、2030年ぐらいにはもう得られるのではないかと。一般の方にもインタビューしましたが、死の看取りをAIにされてもいいけれど、という人が結構多くて、ご家族や周りの人に聞いてみてください。ただ、制度と人の心という問題はありまして、死を特別視しない文化をつくるとか、安楽死

に關係する法律を整備するなどは10年仕事なのかなと思います。

あと動画だけお見せして終わりにします(図6)。どのような世界観かということで4分ぐらいの動画を。今はいろいろな学会でこういう話を聞いたがつて、わりといろいろなところや製薬会社に呼ばれたり、そういうことが増えています。

(動画視聴)

すみません、長いものを見ていただいて。あとはまとめて終了したいと思います。

先ほど申し上げたように、AIは2024年を変曲点として既に新たな時期に入っている。Early-phase Singularityと私は呼んでいます。医療におけるAIは今後、急速かつほぼ全領域に拡がっていく。1例として、「死」にまつわる医療が例外にならない萌芽もあると考えています。以上です。ありがとうございました。

堤 ありがとうございます。特に最後の動画はグッとくるものがありましたけれども、何か先生方からご意見やご質問等ございませんか。

柳原 先生、途中でお示しになった、おばあちゃんと医師が話していて、下に動物がいる。

奥 4者モデル。

柳原 先生がおっしゃったように、患者さんはたくさん話したくて、医師は短くしたいということがあって、それはとても大事なところだと思います。

これだと患者さんは、ご自分が満足するんですかね。誰に話すのでしょうか。

奥 ヒツジに話す。まず、おうちでいっぱい話して

から来るのです。

柳原 ヒツジに話すんですね。そこは人間的なアバターにするとか、そういうのは将来的に何かお考えでしょうか。

奥 アバターにすればいいと思うのです。例えば、よく設定して孫娘にするというのがあると思いますが、映像、絵をつくるにあたってジェンダーの問題が結構難しいので、女性だと女性と決めつける、とにかく孫娘と決めつけると批判されるかもしれないので、ということで動物にしている。アバターはあり得ると思います。

柳原 あそこは患者さんの満足度が上がると、随分AIの診療も質が上がるかなと思います。

奥 そうだと思います。なので、ご本人がよければ孫娘でもいいですし、娘さんでも息子でもいい。でも世の中には孫娘と仲が悪い人もいるからということです。

ひとつだけ言い忘れてましたが、この映像はフルAIで、僕はシナリオを提供しただけで何もやっていません。あれをAIが作れてしまう。全く何もやっていないです。そういうことができる会社には頼みましたが、会社の人がAIをディレクションするだけで、あれができています。最初は脚本家と監督のご夫婦の知り合いがいて頼もうと思ったのですが、二つ言われたことがあります。一つは1年半かかると。もう一つは、原作者は何も口出させません。映画というのは脚本家と監督のもので。原作者のものではないので、奥先生、アイデアとお金は出していいけれど、中身には一切何も口出



1分バージョン



5分バージョン



図6

せませんからと。

それは困るなと思っていたら、AIの会社の会長さんが、じゃあ俺が作ろうかと作ってくれた。1分の動画と5分の動画があって、1分のは一晩でできました。5分のものも1週間ぐらいできちゃう。どのぐらいお金がかかったかは、お金を出していないから分からないですが、たぶん5分の動画で20、30万円ぐらいじゃないかなという感じはします。これのフルコンテンツ、2時間のものをクラウドファンディングか何かで作ろうと目論んでいます。

下澤 実臨床をやられている先生方のアンケートの結果からは、二つのターゲットがあって、一つは医療面の、もう一つは経費削減のところが重要視されていました。現在オンライン診療と対面診療では医療費に差がつけられています。しかし、AIの使用の有無にかかわらず医療費はフラットですよ。でも今後こういうAIのようなものが進んでいった場合に、そこは格差をつけていくことが可能性としてあるのでしょうか。

奥 私はYESだと思います。その意味では、先生がおっしゃるように、今の医療の仕組みがそもそも不十分どころがいっぱいあります。ご存じのように医師法ははるか昔、70年たっています。国民皆保険が始まったのも1961年なので60年以上たっている。コンピュータを前提としなくて作られているので、そこは前提として医療制度も保険外診療も変えていかないといけないと思います。

下澤 実際、AIのいろいろなアプリが開発されている台湾やアメリカではいかがでしょうか？アメリカの場合は自由診療が認められている部分がありますから、もしかしたら値段を変えていたり、保険によっては使える人と使えない人がいるかもしれません。海外は差をつけているんですか。

奥 差はつけているようです。差はつけているようなのと、あとは個人情報に関わる場所の制約が相当弱い。日本だけが突出して世界で強い。日本でもその問題さえなくなればそうなると思います。

診療報酬のことばかり考えていると、おっしゃるように対面とどうという話になるのですが、コストがすごく下がれば自由診療とうまく組み込んで、ほどほどの値段のものを診療外で使う。それが前さばき的に、診療報酬で手当しなければいけない部分に到達する前に、前さばきのところで医療の質が上が

ると思う。そういうモデルを厚労省や内閣府が先導しなければいけない時期にきているのかなと。

下澤 そうすると予想として、先生は今、放射線科のことをおっしゃっていましたが、明らかにAIのほうが人を超えてしまった場合は、AIのほうがコスト、診療報酬が高くなってしまいます。それとも人間が読んだほうが高くなるのでしょうか？

奥 今の医師法のままだと、人間の医師の診療報酬が高止まりというか、AIから超えられることはないんじゃないでしょうかね。今は判子は人間の医師が押さないといけないモデルなので。

下澤 そうですよ。でもそれが将来的にAIでいいということになればどうでしょうか？

奥 そこはですから、医師法を変えるという議論になってきます。

下澤 現在そういう方向にあるんですか。

奥 いや、全く議論されていないと思います。議論してほしいです。

下澤 そうですよ。先生としてはどちらがいいと思いますか。

奥 AIを含むしかないと思いますね。

下澤 AIのほうが高いほうがいいのかと思いますか。

奥 そうは思ってないです。結局たくさん使われるとコストが安くなるので、アメリカの自動運転がとても参考になると思って著書にはいろいろ書いているのですが、自動運転があまり日本では報道されていないけれど、死亡事故も結構たくさん起きている。だけれども結局、経済合理性という観点で自動運転が進んでいっている。

医学にまつわるところは先生方も皆さんご存じのように妙な感情論があって、命は地球より重いみたいな、物理学の法則を無視したような話があるので、そこから変えていかないといけないとは思いますが、個人の能力を超えるので、そこは国として誘導すべきかと。

下澤 もうひとつ、経済の問題ではなくて医療面で、医者や医療従事者が自分の仕事を少しセーブできるのではないかということについてお伺いしたいです。昔からアメリカだったら手術記録はディクテーションでした。自分で書くなんでやっていなかった。これからはそれをただAIがやってくれているだけ。オペレーターに自分で電話でしゃべらなくていい分、医者としてはやっていることは一緒だと思うん

です。AIがやるか、オペレーターがやってくれるかだけの違いだと思いますけれど、そんなにAIを使うことによって、医者なり医療従事者の仕事は減るんですか。

奥 いったんは減ると思います。雑務というか、人数的にマジョリティな仕事が減ると思います。ただその分、もっとガイドラインを超えた診療の選択とか、日常的な言葉で恐縮ですが、寄り添いの部分とか、そういう部分に人間の医師が深くコミットできるように。おっしゃるようにディクテーション的な、AIがやっても人間がやってもいいところはほとんど全部解放される。

下澤 僕もディクテーションやっていたんですけど、最後は校閲確認が結構面倒くさいですよ。

奥 そうなんですけれど、プルーフリーディングみたいなところこそ、人間の手もほとんど要らなくなってくる。今だってAIとAIでクロスチェックさせることはできるわけですから。AにやらせてBに批判的に見てねとやるだけで相当クオリティが上がる。

下澤 先生がおっしゃるとおり、ディクテーションのシステムが日本でもたぶん使えるようになるという。実際にしゃべっている発音やアクセントの違いはどこまで正確に読んでくれるのでしょうか？

奥 ほとんど問題は起きないです。

下澤 例えばスマホの音声入力は僕は全然駄目なんですけれど。僕のしゃべりだと、読み込んでくれないんです。

奥 読んでもらえないですか。

下澤 読んでもらえないですね。かなり。

奥 そこはAIの技術者と僕はよく話しますが、少なくとも先生の音声を登録すればちゃんとできると皆さん言うんですけど、そこももう、あと1年ぐらいで話者特定は必要がなくなると思います。

先ほど介護の現場の話をしました。介護の現場は東南アジアの方が多くて、いろいろな言語をしゃべるわけです。でもそれでも全然ブレは困らない状態。介護システムは実際に導入が始まっているので、言語はあまり気にしなくていい。東北弁でも大丈夫。

下澤 僕、実際に一番使っているのは翻訳アプリのGoogle翻訳なんですけれど、あれの発音が、英語などはいいですけれど、日本語があれば意外と悪いです。漢字が間違っていたりとか。だからどれぐ

らいまで精度が上がっているのか、プルーフリーディングはどれぐらいまで必要なのかなど。

奥 同音異義語の問題の難しさというのが結構あって、漢字の微妙な選択。でもそこは、私もすごく気になるんですけど、医学的なコンテキストで言うと、極論ですが、誤字があっても、それを後で他の文章と整合して振り返ると、これは誤字だよねと分かるので、それほど神経質になることではないのかなというのが今の暫定的な結論です。

下澤 僕も期待はしています。

奥 もう1年、2年でびっくりするぐらい精度がよくなってくる。

下澤 でしょうね。3年前のGoogle翻訳と今では全然違うので、それは確かにそうかもしれないですけどね。でも40年前にディクテーションをやっていた頃に比べたら、格段に楽になりました。

奥 あの頃はトランスクリバラーがいらして。

下澤 大変でしたけれどね。

大西 たぶんAIが奥先生がおっしゃるようになるのは間違いないと僕も思っていますが、この静かなSingularityというのは現在は先端レベルでそういうことが起こっているということなのかなと。今はまだ人工知能を動かすのにかなりエネルギーが要するという話で、そこの兼ね合いや、あとは実際に装置を全診療所が同じように持てるようになるのかとか、そのあたりが気になる場所ですね。

私としては結構格差が出てくるのかな、AIを使えるところはよくて、使えないところは人間がそのままやったりと、そういうのが残るのかなと思ったりするのですが。

奥 2点ですけど、Singularityの話に関しては大西先生がおっしゃるように、今、先端の部分ではそうだけれど、世の中に広く浸透するという意味では5年ぐらいのタイムラグがあるだろうと。

大西 でもその程度のタイムラグということですか。

奥 そのレベルだと思います。5年と私が申し上げる理由の一部でもあります。今おっしゃった経済格差があって、一部のポピュレーションには使えないのではないかという話に関しては、コストが、その生活レベルというとあれですが、それを超えて下がってしまうと、もうそっちを使うほうがよくなってしまふので、そうすると急にダーッとそちらに行くんですよね。なので、意外と早いと思います。



大西 宏明 先生

大西 テレビや洗濯機が普及したのと同じようにですね。

奥 収入レベルにかかわらず、テレビやスマホを持っているのと同じ。

大西 その程度のレベルでしかエネルギーは使わないので、十分実現可能になってくるイメージですかね。

奥 そんなイメージだと思います。

大西 死への踏み込みに関しては、おっしゃるとおり現在の日本は死をかなり特別視していますが、日本人の伝統的な死生観は必ずしもそういうわけではないので、先生がおっしゃったように、トータル的に見てそのほうがいろいろなことがあれば、たぶんそのようになってくるのかなと。

奥 ならないとしょうがないのかな。

大西 介護の問題も大きくて、今はそこに使われる人のコストが大きすぎる。そこをAIでやってもらえるようになると、解決できる部分もかなり大きいように思います。先生は研究として、まさに最もやるべきところをやっているから素晴らしいなと思いました。ありがとうございました。

下澤 これのデータのバックアップはどうするんですか。例えば放射線科の画像のバックアップはどこかに置いておきますけれど、震災のときに電力が落ちるとアクセスできなくなりますよね。

奥 一時的にアクセスできなくなることは、2個目、3個目のコピーにどれだけお金をかけるかということになるので、そういう意味では、問題は災害時や事故時に起き続けると思います。

そもそも今の Amazon や Google などが預かって

いる情報を見ても、少なくとも1個とか2個しかコピーがないという感じでは全然ないので、ストレージを48時間に1回とか全部取って、それを2か月でも1か月でも入れ替える。そういう意味でいうと2週間×4ぐらいのバックアップは常にやって、どこまででも遡れるようになっている。Ready to use に関してはおっしゃるような今は使えないということはあるかもしれませんが、データが損なわれることは今後人類では二度と起きないと思います。

下澤 例えばそういったときに、学習したことを忘れてしまうことはないんですか。

奥 それはないです。ないですと僕が断言するのは、そもそも学習した知識というのは、情報と、その情報からどうやって知識をつくったかというプロセスが定義されているものであって、生産されたものがストレージされているわけではない。もともとのデータがあって、それにどういうプロセスで推論したかというアルゴリズムが存在する。その組み合わせなので、結局再現性があるわけです。なので、停電だから知識が失われることは起きないです。

そういう話をするすると疑問として出てくるのが、2回目、3回目をやると違う結果が出るんじゃないかと思うと思いますが、それは10回とか100回やればいいです。多数決を取って結局やり直すとこの結論が多いと。今のAIは全部それです。全部多数決でやっている。囲碁が人間チャンピオンを破ったのも、最後まで、100万回打ちちゃうわけです。100万回打って一番勝率が高い手を打つという、すごすぎるんです。

下澤 死に関して言うと、例えば死生観は、当然宗教のバックグラウンドによってかなり違ってくるので、どこで使ったロボットかによってもユーザー側が選ばなければいけなくなってくるのではないのでしょうか？

奥 話が雑駁になるのを避けるために入れてなかったのですが、安楽死のマシンでサルコというのがありますが、安楽死のマシンはオーストラリアで開発されて、ヨーロッパ、スイスなどでそれを使うことが法的に認められたりしたプロセスがあったのですが、その創設者がなぜか不慮の死で死んでしまったという、結構ドロドロした世界で。

2年前にも死に関する法律が、詳細は今述べませんけれど、フランスで法ができたときにも、いつ死

を認定するかということで紛糾して裁判所の判断が分かれたりして、そのときにフランスにいたので、フランス人の人と結構議論したのですが、死生観が日本人だけ遅れていて、ヨーロッパの人はクリスチャンだから大丈夫という感じではない。彼らは彼らでちゃんと悩む。

なので、あまり日本は死生観が発達していないからというディスアドバンテージがあるという感じでさえないというのが今の私の感覚です。

堤 質問も尽きないと思いますが、時間のほうもあるので、また後ほど別の席でということで、ありがとうございました。

それでは次に長崎大学の柳原先生からお話をいただければと思います。ご専門の領域である、微生物検査領域におけるAIの現状や問題点、今後の展望についてお話をいただければと思います。どうぞよろしくお願いたします。

検査・診断領域におけるAIの現状と未来 —微生物検査—

柳原 私が考える微生物検査・診断領域のAIと現状と未来について微生物検査を中心にお話します。

微生物検査の進歩というところで、培養や染色などあったので、1990年ぐらいから遺伝子検査は始まって、今ずっと進歩している。それと、質量分析



柳原 克紀 先生

装置というのがここ5年から10年ぐらいで、AIは最近入ってくるだろうという年表です(図7)。

私が委員長を務める感染症学会AI委員会で人工知能に関するステートメントを作成しました。こういうのをつくって一回考えましょうということで、毎年感染症学会ではシンポジウムを行って情報共有をしています。

まずはグラム染色におけるAIです。グラム染色というのは画像なので比較的進んでいます。2018年辺りから検討されています。グラム染色画像を取り込んで分類しましょうということで、かなり精度が高いのがもう7年前に出ています。

多くの画像を入力することで優れたAIができま

微生物検査の進歩

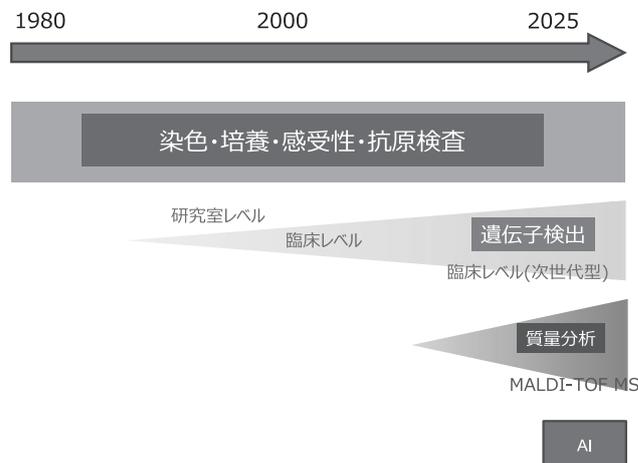


図7

す。ROC 曲線を見てもかなりいいものができています。

わが国の現状ですが、日本の会社が機器を開発し、研究を進めており、長崎大学も協力しています。染色、画像撮影、診断まで自動化されています。

臨床検査技師と同様に染めることができ、撮影と診断までも自動でやってくれるので、かなり省力化されています。

グラム染色のワークフロー、自動化、AI サポートです。自動化することで、より削減しようということで、自動化+ AI というのが微生物検査等では進むかなと思っています (図8)。

薬剤耐性菌における AI 検査も開発されています。今から MRSA の話をしますが、わが国では黄色ブドウ球菌のメチシリン耐性が多いので、これをどうしようかということです。

耐性菌は随分嫌がられていて、2011 年だと、たぶん技術が不十分だったこともあって、少しくまいていませんでした。

3 年前に出た MRSA の同定で、先ほどもありましたが台湾は進んでいて、2 万株を使って MSSA と MRSA を分けて、質量分析のパターンを使って判読しようということが出ています。

AUC を見ると 0.9 ですので、感度も特異度も高いことになります。

論文では、目印となるものを決めて MRSA の特殊なパターンと特定しています。そして、三次元解析でもだいたい妥当だろうということによって論文になっています。

また、私たちと台湾の三軍ホスピタル、Tri-Service General Hospital、防衛医科大学と一緒に、MRSA の毒素が強いものを判別することでできないかと研究を行っています。約 34 万 5,000 の検体を用いて、MRSA の遺伝子型別が AUC 0.9 で、PVL というのが毒素が強い株ですが、それも 0.85 ということで、報告しています。

最近の私たちの微生物の領域でかなりトピックスになった論文がありますが、培地の培養プレートに関するもので、培地は色も形も全然違うのですが、深層学習と階層的 AI を使って混合感染や単一感染などを判別することができないかと。今までは慣れた検査技師が、これは感染菌だろうとか、常在菌だろうとやっていたのですが、それを機械でやるということです。

細かいデータはお示しませんが、コロニーを取って、それを読んで実際のデータと合わせてい

微生物自動染色分析装置 Mycrium® のワークフロー 対応材料種：血液・尿 ※その他材料も現在開発中

簡便な操作と自動染色から AI 判定サポートにより夜間休日のグラム染色結果報告が可能になります



株式会社 GramEye 様よりご提供

図 8

(図 8 は巻末カラーで掲載しています)

て、それでもかなり膨大な作業ですが、データとしてはかなり使えそうだと思います。これはおそらく臨床検査技師の労力削減になると思います。随分削減になって、この辺りにすればいいよとか分かってくるんです。自動塗布装置と絡めると、ますます労力削減には有用かなということで、これは注目された論文の一つです。

抗菌薬のことを少し調べてみると、スクリーニングする、作用機序を推定する、AI×MALDIによる耐性機構、ゲノム解析で新規耐性遺伝子など、おそらくいろいろなデータを取って、AIにかけて、抗菌薬を作ることが将来できるだろうということです。

先ほど堤先生からご紹介がありました、nodocaというのは日本初のAI新医療機器です(図9)。これは写真で喉を見るだけでなく、最初に問診入力があります。これが結構大事で、例えば昨日から喉が痛いなどの臨床情報をいれることで、精度が上がっているんです。ですから、病歴、経過および咽頭の写真の情報をAIで分析して診断するものです。インフルエンザ濾胞という特徴的な所見が診断に役立つようです。

免疫クロマト法というのは検体の採取があって、そこで感染するリスクが出てくるので、撮影した写真を用いると、そのリスクが低いだろうということです。

感度はかなりいいです。感度と特異度が70と85と、免疫クロマト法に近いものなので、それでは保険点数も取って検査法として認められています。

今後の感染症診療とAIですが、ご紹介したLaboratory&imaging diagnostics、Clinical decision supportということで、病歴などを入れて、いろいろな感染症を拾うもの、感染のPublic health & surveillance、あとは先ほどご紹介した創薬とかその辺りです。

問題点は法規制・品質保証です。特にAIだと検査に求められる外部精度管理が難しく、それをどうするのか。倫理・プライバシーはAI全体の問題ですし、患者データの二次利用、AI学習用データ共有もAI全体の問題です。責任所在もAI全体の問題ですが、検査でもAIが言ったから診療して間違ってしまうかということもあります。

まとめますと、微生物検査のAIは画像やスペクトル、今日はお話ししませんでした。NGSのデータももっと使おうというのが論文としてはあります。あとは同定と薬剤管理AST、アウトブレイク検知、AMR監視は間違いなく大事だろうということです。

データバイアスや説明可能性。結局私も思ったのは、説明可能性、説明できないですよね。これはなん

インフルエンザの診断（日本初のAI新医療機器、nodoca）



Okiyama S et al. J Med Internet Res. 2022 Dec 23;24(12):e38751. doi: 10.2196/38751.

アイリス株式会社様よりご提供

図9

でブドウ球菌なのと言われても、MRSA なんて、と言われても、感受性の結果を取っていないので説明できない。これはどうするのか。あとは規制・倫理。

ただ、ガバナンスを使いながら、先ほども奥先生からありましたように、AI が進歩して使うのは間違いないので、ガバナンスを作って、役割を決めて、人間の監視の下で、安全かつ効果的に検査領域で使うべきだろうというのが私の結論です。ありがとうございました。

堤 何かご質問、ご意見はありますか。

私から一つ。MRSA と MSSA で分ける論文を読んでいるのですが、質量分析においてピークが MRSA だけ見られています、あのピークは何を意味するのでしょうか。

柳原 本論文の MRSA 検出機序は、MALDI-TOF MS で得られるタンパク質スペクトルの特徴を機械学習で解析し、階層的分類によって MRSA を同定する AI ベースの間接検出法です。

堤 PVL など毒素を作るタイプだと、たぶんこういうピークが出て区別がつくのかなと思いますが、なるほどそういうことなのですね。分かりました。

もう一つ後半の nodoca の話ですが、最近確かコロナ用も作られて、まだ診療にはどうかという感じですけど、インフルエンザはどの時期も、たぶんそんなに臨床症状や喉の所見もあまり変わらない。ただコロナの場合は症状も変異が進んで変わったり、喉の所見は変わるか分からないですが、その辺の変異株に対応した形できちんと診断ができるのか、不思議に思うのですが。

柳原 それは臨床検査医学会の感染症委員会で審議して、結論から言うと、nodoca はコロナに関してはインフルエンザに比べると精度は高くないです。ですから、陰性に出ると結構言えるんですが、陰性になる可能性が全体の 20% しかないんです。

堤 だいぶ広く引っかけているのでしょうか。

柳原 引っかけているんですよ。ですから、こちらの言い分としては、インフルエンザを診断するとき撮影した写真を使って、インフルエンザがない場合にコロナかどうかは判定できますよということなんです。感度と特異度が陽性になるときも 20% ぐらい、陰性も 20% ぐらい。あとの 60、70% が分からないんです。ですから、いわゆるボーダーがほとんど占めるんです。

だから彼らが言うには、インフルエンザではない方で、コロナと診断できれば使えます、そこで保険を取りますよということで申請しているのですが、もうひと工夫望まれるところですね。方向としては大事なのかなと思います。

堤 確かにそうですね。インフルでなかった場合にコロナが濃厚という形でした。分かりました。

奥 先ほど説明可能性というお話が出てきて、そこは私も関心があるところなので、先生のお考えとしてどの程度の説明可能性を求めるのかと。

ご承知のように 1990 年代ぐらいからエビデンスということが言われるようになって、EBM を基に医学は発達してきたのですけれども、AI の手法が出てくるようになって、人間には説明してあげられないけどこうなんだということがすごく増えてきた。さっき私が例に出した囲碁のところでも、NHK の解説者のプロの先生たちが、これは分からないけれど正しいみたいですねという話になっている。

そこを先生方の学会としてはどうする、これは医学会全体も同じ課題があるのですが、どこまで説明責任を求めるのかということについてお考えを伺えますか。

柳原 感染症学会のステートメントはそこまでまだ踏み込んでなくて、今回は論文を引くと説明可能性のことが出ていたので入れました。例えば生化学検査だと、化学反応等としますけれど、そこをわざわざ臨床の医者に、ここは化学反応をしていますから、こうなっていますよという説明はあまりしないので、将来的にはその辺りが納得していただければと思いますが、現時点では難しいので、将来的に AI が普及してくると可能かなと思っています。それも過渡期にあるのだらうと思います。

奥 ブラックボックスという言葉がよく使われると思いますが、ブラックボックス化に関しては、AI の世界、僕ではないもっと工学よりの人たちの間でも結構問題になっていて、ブラックボックスであるから駄目という話をし続けると合理性が担保できないので、ブラックボックスをどう認めるのか、認めないかということのをわれわれが、今後の数年間で直面していくのかなと。

数学の定義もここまでできているけれど、この数学で、この n を一般化して解いてくださいとか、その辺を AI ができるようになってきて、今年の 2 か

月ぐらい前に結構典型的な例が出たのですが、そうなるブラックボックス、人間が分からないからと言ってられないということが大きな問題としてあるのかなと。答えのない解説ですが。

大西 おっしゃるとおりかと思います。もともと人間が全部分かっているわけではなく、ブラックボックスとか偉そうに言っているけれど、実はほとんどブラックボックスでやっている。疾患の何とかの三徴というのは、何の関係もない症状を3つ並べて、それからこの病気だろうみたいな、単にそう言っているだけなので、実は何もそこは分かってないですよ。例えば手足口病だって、何で手と足と口に発疹出るのか全然分かってないですが、それでも診断はしている。それでローコストで、ある程度カテゴリー化して、治療をする・しないも含めて判断している。だから、安全性も含めて経験的に実用可能なものであれば、ブラックボックス問題はあまりこだわらないほうが良いようにも思ったりもするのですけれどね。

最終的には先生がおっしゃったように、データとその後のアプローチの仕方が残っていれば、もしかしたら後でそこから何かヒントが出てくるかもしれませんけれどね。

奥 いつか分かるかもしれない。

大西 ええ、いつか分かるかもしれないと思います。今の時点ではそれが全部分からないと使えないということにはしなくていいように個人的には思います。

下澤 検査のほうから細かい話で申し訳ないですが、外部精度管理の話がありました。生理機能検査でもそうですが、外部精度管理は結構難しいですよ。例えば心電図読みをどうするか。AIに学習させるときに精度管理ができて心電図データを読ませないといけないのでないかという話がありますが、これはどうしたらいいですか。

柳原 今の、例えば微生物であれば、菌株を多くして特定させますよというようなことになるとは思います。そういうことはできませんので、AIをむしろ外部で管理するとか、あるいはスペクトルのパターンを送って、それを特定できるのかということをするとか。たぶん精度管理自体が違う方法に進むだろうと思います。

下澤 今回のコロナもアウトブレイクで、全く今ま

でわれわれがデータを持っていないウイルスが出るわけじゃないですか。スペクトラムが全然違うものが出てきたときに、そういう精度管理でどのように引っかけしていくのかとか。AIの使い方、今先生が示されたのは、分かっているものに対してはできる。分かってないものが出たときにはどうするのかというところは、どのように。

柳原 そこも例えば今まで分かっているのがあって、分かっていないものが出てきたら、分かっていないと。ということは未知のウイルスですよ。最初はそのレベルで、それをもっと増やして行って、新しいウイルスと特定するものを使うんですかね。

大西 精度管理は学会（日本臨床検査医学会）としては非常に重要なところですよ。心電図はどうなんでしょう。画像のパターンで出るので、AIが練習しやすいのかなと。

下澤 実際は、電位が不正確だったり、紙送りのスピードが悪かったりあるようです。

大西 それを何回もとれば相殺されて、結局ある程度平均的なものが出るのかなと。逆に血液検査の方は測定試薬や機器のバリエーションが大きすぎてなかなか難しいなという気もするですよ。

柳原 心電図は随分前からありますよね。あれはこの10年ぐらいで、だいぶよくなっているんですか。

大西 よくなっていますね。

下澤 今度LMIに論文が出ますけれど、結構ブレるみたいです。

柳原 よくなっているけれど、まだブレる。

下澤 しかも外部精度管理のやり方が、スパイクを入れてちゃんとそれが出るかだけなんです。だから僕らが普段検体でやっている精度管理と全然違います。人工的な心電図を作るようなことはしていないです。エコーもそうで、深度と輝度が一定かどうか。

そういう意味では先ほどのグラム染色も、機械によってかなり染色度が違いますよね。それからその部屋の温度などで違ってきたときに、染色がどれぐらい精度よく読めているのかなと気になりました。

大西 今心電図のデータを使った色々な診断基準を作っていますよね。あれも人為的なもので、例えばAIなどで病気と心臓の電気信号との対応型ができるようになれば、逆にそちらのほうが正しいというか。とり方はもちろんある程度そろえないと

いけないけれど。

もともと揺れのあるデータをたくさん集めたときに、それをどう解釈するか。精度管理と違うかもしれないけれど。そういうのはAIは得意なのかなと思ったりしますね。

奥 それについてですが、心電図はモダリティとして典型的に先行きが不安。というのは、そもそも12誘導をとって、どこに付けてということが、そもそも情報を制限して、要は人間が理解可能な形に落とし込む技術ですよ。おそらく、本当は今でさえそれをやめてしまって、とり方からAIにやらせたほうが絶対まともなはずなんです。でもその議論を今はなかなかできないので、もう少しいろいろなことが落ち着いてから、ところで心電図はどうしようかという話になるのかなと僕は思っています。

下澤 昔からある技術ほどそうなるんだと思います。

奥 おっしゃるとおりです。昔からある技術ほど、先に情報を制限して、そこからスタートしているので、方法論から変えていくということ。

下澤 それこそブラックボックスの中をぐしゃぐしゃとやっているようです。心電図検査の所見を客観的、統一的に表現するために、アメリカのミネソタ大学で発案された所見のコード体系って偉大だと思いますけれど、難しいですね。

奥 結局ダウングレードして、先ほど先生がおっしゃったような三徴の話もそうですけれど、ダウングレードしないと人間の手に負えないので、医学がそうやってダウングレードすることによって発達してきた歴史が数百年あるのだけど、そこを捨ててはいけなくなるわけです。

堤 AIが検査そのものを変えていくという感じですね。他はよろしいですか。ありがとうございます。それではここからは大西先生に司会をバトンタッチいたします。

大西先生、よろしくをお願いします。

大西 次は国際医療福祉大学臨床検査医学教室の下澤達雄先生から、検体検査領域におけるAIの現状と未来についてご説明いただきます。特に、先生がこれまで携わってこられた、AIを活用した尿沈渣の識別や、顔認証による検査値の推定なども含めてお話しいただければと思います。また、既にお話を

いただきましたけれど、生理検査も含めた検査領域におけるAIの利用の現状や問題点、今後の展望についてお話をよろしくお願いたします。

検体検査領域におけるAIの現状と未来について

下澤 皆さんが素晴らしいスライドを用意されていて、僕はすごくいい加減なスライドしか用意してなくて申し訳ございません。先ほど奥先生のほうからご紹介がありましたが、僕のバックグラウンドは昭和63年に筑波大学を卒業して、東大の内科で研修を始めて、2017年まで東大にお世話になっていました。基本は高血圧が専門で、2000年から臨床検査に移って、高血圧なのでホルモンの分析とか、尿を見て腎臓の機能を見るところを検査部、あとは遺伝子の診断を少しやらせていただいていた。内科医としてよりも臨床検査医としてのほうがキャリア的に長いのですが、圧倒的に自分がやっていることは高血圧です。

高血圧学会も今年新しいガイドラインを出して、130/80 mmHgにしようという話をしたのですが、それを宣伝するためにYouTubeを作っているんです。奥先生からはフルAIで作られたムービーがご紹介されましたが、作成しているYouTubeもフルAIです。しかも業者に頼んでなくて、スクリプトは僕らが書きました。書き方にもよりますが、30分から1時間で1分ぐらいの動画がすぐできてくるんですね。今は音を鳴らしていませんが、曲も全部AIが作っています。

(動画視聴)

こういうのが簡単に作れてしまう時代になっていて、私はAIは完全に利用する側でありまして、理論や倫理といったようなことは全く考えたことがありません。

そういう具合で、利用する側としてどのように使っていくかということで、今、私がやっていることをご紹介させていただきます。

もともとファブリー病という病気があるんですね。希少疾患で、小児で発症する場合と大人で発症する場合がありますが、異常なタンパクが蓄積することによる疾患です。酵素活性がなくなることでX染色体連鎖性遺伝なのですが、その酵素を補充し



下澤 達雄 先生

てあげることによって発症しなくなる、あるいは臓器障害を起さなくなることが分かってきて、早期発見、早期治療ができるようになってきた疾患です。もちろん酵素活性を新生児マススクリーニングでやればいいのですが、費用が掛かるため一部の県でのみ新生児スクリーニングの中にファブリーを入れているところがあります。

その診断でわりと簡単にできるのは、尿中に出てくるマルベリー小体とって、脂肪の塊のようなものが細胞にくっついている、こういうマルベリー小体、マルベリー細胞ともいうことですが、これが尿中に見つかり、かなりの確率でファブリー病だろうといわれています。マルベリー小体、細胞が陽性であれば、酵素活性を確認し、さらに遺伝子変異を確認します。シャペロン療法が使えるか、使えないかは遺伝子変異を見ないと分からないのですが、そういうことができるようになりました。そこで新生児スクリーニングをしようということが広がってきました。

ただ、このマルベリー小体というのは、尿沈渣を全視野見て、1個か2個見つければいいほうで、技師さんたちは大変なんです。その目を持って見ないと、これはネフローゼのときに見えてくる、マルタ十字というのが見える卵円形脂肪体と見間違えてしまったり、あるいは訳の分からない上皮細胞だと見逃されてしまうこともある。

聖マリアンナ医科大学の臨床検査の教授の右田王介先生は成育医療研究センターとの共同研究で作成された、スクリプトを報告されています。これが尿沈渣の細胞成分でやったAIの最初の論文だと思います。世界中でもこれが最初の論文で、使っている

ところは結構あると聞いてはいますが、無料で公開されているので、使おうと思えばいくらでも使える。こういうのがあります。

ファブリー病のサンプルをいくつか学習させた上で、それ以外の細胞も学習させることによって、AUCが十分に大きいROCカーブを得ることができています。

URL、https://github.com/huryu/DL_Fabryに全部コードは公開されていて、僕らも使っています。この問題はどこにあるかという、未染色、すなわちS染色で染めなくても見つけれられるのが売りなのですが、実際には染めないと判別が難しいところがあったり、前に最近のグラム染色の染め方の標準化が難しいという話がありましたけれど、S染色の標準化はちょっと難しいところがあります。

それから全視野を見るのは人間の手で見なければいけないので、まだ時間がかかるなというところでは。

右田先生といろいろな話をされていて、彼がちょうど筑波大から聖マリアンナ大に来たので、尿沈渣をAIで検査できたらいいよねと話をしていました。ファブリーは僕らも興味があって技師教育をやっていたのですが、彼らがこれを出してくれたので、次にやるものとしてわれわれは尿沈渣をもう少し広くAIで診断できないかを考え始めました。というのは尿沈渣はpoor man's biopsyとも呼ばれていて、腎臓の中で何が起きているか生検をしないでも、ある程度沈渣細胞を見ることによって診断できます。

沈渣の歴史は非常に古くて、ベルツ先生の診断学の教科書にも出ているんです。あの頃の顕微鏡で、すでに事細かに書いているのですが、やはり、その中で見逃されている細胞がいっぱいあって、例えば僕らは丸細胞というのを見つけて、尿細管の再生を見ているマーカーとして使えるのではないかというのを報告しています。また、ポドサイト、糸球体上皮細胞というのがあって、これが落ちてきている人たちは腎機能が悪くなる予測因子になるということが知られていたので、このポドサイトを診断できるAIを、ファブリーと同じようなものでできないかと考えて、今、大学院生の横山千恵さんがやっています。

これは何が大変だったかという、尿沈渣を見るとたくさんいろいろな細胞が落ちてきているのです

が、1個1個アノテーションして、これは何細胞、何細胞と名前を付けていって、その中にポドサイトはどれぐらいいてというのを覚えさせる学習が大変でした。現在1,500枚ぐらい読ませました(図10)。

読ませて、とりあえずスクリプトができて、今バリデーションを取っているところですが、このバリデーションの取り方も他のサンプルで読ませるといってもそうですが、他の先生たちが、他の検査技師たちが見て付けた細胞の名前と同じかどうかというバリデーションも取らなくてはいけなくて、今それを沈渣の専門家の先生たちと一緒にやっています。

横山さんの博士課程の研究なので、3年間で仕上げなくてはいけなくて、今2年目なのですが、今バリデーションのところになっています。あとは論文を書いて終わりにしようと思っています。だからこれはまだ論文になっていないので、紙上に画像を出せないのですみません。こういう内容をやったりしています。

もう一つは先ほどあった顔認証を用いた血液検査データ取得方法の開発で、これは今年の臨床検査医学会の一般演題で発表させていただいたのですが、SFMという会社と一緒にやっています。何

をやっているかということ、なるべく患者さんたちが痛くないように、楽に検査ができないか。採血しないでいろいろなことが分かったらいいよねということで、技術革新の最前線では皮膚や爪の色でSpO₂を測ることができるようになってはいますが、赤外線体温を測ったり、超音波によってリアルタイムのモニタリングというのがあります。スマートフォンによる画像解析で、自撮りをしたことによるいろいろなデータがだいぶ分かってきました。

採血は痛いし、高齢者や小児だと大変です。それから日本では在宅医療が問題になっていて、臨床検査を在宅医療でどうやってやるか。ポータブルのエコーの機械ができて、確かに在宅医療でもできるのですが、できる技師の数が限られていますし、医者でもエコーをとれる人が限られているので、やはり難しいといわれています。それに対して検体検査だったら比較的簡単に大量にできるのですが、採血した後にサンプルをどうやって保存しておくか。例えば放置しておいて冷蔵保存にするとあっという間に溶血します。それによってデータに誤差が生じる問題があるので、病院ですぐ分析するのは在宅医療の場合にはだいぶ違ってるところから、それを克服できるのではないかと。画像診断でできないかと

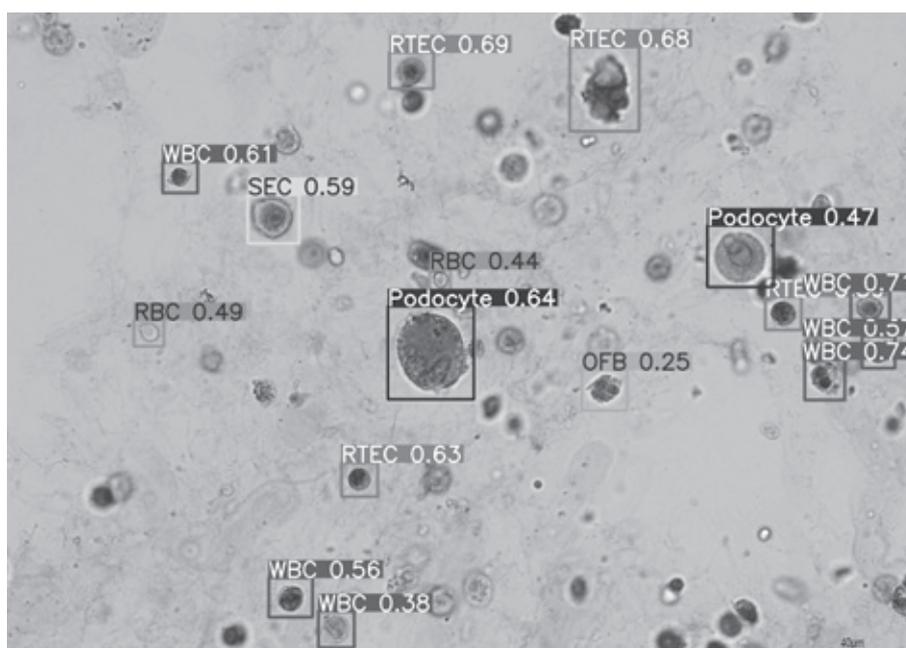


図 10

(図10は巻末カラーで掲載しています)

いうことを考えています。

医療現場で例えば血清を取ろうと思ったら凝固させて10分待って、遠心して10分待って、それから分析が始まるので、結果が返るまで、どうしても1時間ぐらいになってしまう。それを、今僕らがやっているのは30秒の撮像で測れるようになってきました。そうするとスピードアップにもなるし、いいのではないかと考えています。

そして、採血だと例えば医療廃棄物。捨てるのに多額の費用がかかっていますが、写真だったらそういうのがかからないし、消耗品も減り、試薬代も浮くということで、栄研さんは困るかもしれないですが、医療コストの削減につながる。

あとは病院に来ないでも遠隔診療でできるようになってきて、患者さんたちの時間を取らないで済む、QOLの向上にもつながるのではないかと。それから技師たちの手間も省けるのでいいのではないかと。ということで、こういうこと進めています。

将来展望は何をやっているかという、貧血や糖尿病など慢性疾患へのアクセス性の向上ということです。世界的に今は競争なのですが、何が問題かというと、アフリカの奥地や南米などでは病院に行くのが大変なんです。ところがみんなスマホを持っているんです。スマホで撮って、アプリでそれが分かれば病院に行かないで済む。先ほどマラリアの話がありました。貧血や糖尿病という生活習慣病もアフリカでは問題になり始めていますし、南米では大きな問題になっているので、こういったようなことに使えたらいいなと期待しています。それから、痛みのない未来への第一歩になるのではないかと。ことを考えています。

どんな機械かということ、実際はスマホのアプリで、2026年にまず英語版を売り出すことになりました。厚労省の認可は取っていません。なので、医療機器ではないですが、日本ではなくて海外で売ろうということを考えています。撮像は普通の家庭用のビデオカメラでもいいですし、iPhoneでもいいということです。アプリはスマホ側に入れておいて、患者さんに30秒撮っていただいて、今僕らはデータをどうしているかという、顔で撮ったデータと僕らが実測した値、精度管理ができたものと比較しようというものです。

今やっている項目はヘモグロビン A1c (HbA1c)、

尿酸、コレステロール、中性脂肪、LDL、Dダイマーが一応測れるのではないかと。ということでやっています。そういった項目をどんどん増やしていこうということをやっています。

今年の8月のデータでは、とりあえずA1cですが、最初に僕のところへ頼んできた会社は、健常者の誤差が19~22%ですごくいいでしょうと言ってきたのですが、先生方ご存じのようにへHbA1cが2割ずれたら臨床的に話にならないわけですよ。そんな絶対駄目ということで、バージョンアップしてくれとずっとやっていて、1,000例近くデータを積み重ねるとだいぶよくなりました。

古いアプリでBland-Altman分析をするとこのような具合で(図11)、確かに6ぐらいのところだとまあまあいいですが、上になってくると全然話にならないというところでした。もちろん僕らはHbA1cを標準化された方法で測っているのですが、異常ヘモグロビン、例えば胎児型のヘモグロビンFが優位の人などが当然いるので、その部分がはじけていないという問題があるのですが、一応こういうデータです。特に5から6ぐらいで検診で引っかかる人たちでもかなりブレているというところでした。これは全然話にならない。彼らはこれで売り出そうと思っていたらしいですが、これは全然駄目だからと話して更なる精度を求めて改良しています。

改良後のアプリもこのような具合で、少し改良しすぎてしまって高めに出るようになっていたりしているところがありました。これがバージョン2になります(図12)。今、バージョン3になっていて、バージョン4を最終的に売り出そうということになっています。バージョン4も8を超える人たちは駄目なんです。だから病気の人たちは基本的に測定できません。健康な人たちのところで、ほぼほぼ安全だと、安心できるデータを提供できればいいかなというところなんです。今後もっと精度を上げていかなければいけないのですが、それは学習させればもっと良くなるだろうと考えています。

というのが現状で、今のところまだ医療の現場で使えるところまではいっていないので医療機器としての販売は考えていなくて、健康のモニタリングでやろうと思っているものをご紹介させていただきました。

あとは、先ほどの心電図やエコーも同じように、

画像なので、特に心エコーなどは技師の腕にかかっているところもありますから、それをできるだけAIで標準化してとった方がいいのではないかと考えています。

それから先ほど奥先生がおっしゃったように、今までの撮影法で本当にいいのかというところがありまして、もっと他の撮り方がいいようなところも出てこないかということを考えています。

そして、もっと組織レベルでいろいろなことが見

られないかということもやっているのですが、動くものは難しいです。静止画はわりとやりやすいのですが、エコーのように動画になると急にハードルが上がってくるといったようなこともあります。

先ほど僕にご質問させていただいたのですが、データのストレージでどれぐらいの容量を用意しておいた方がいいのかということも分からなくて、僕らは10TBぐらいで用意しているのですが、全然足りないのが現状になっています。

旧アプリ HbA1c Bland-Altman 分析 (n=175)

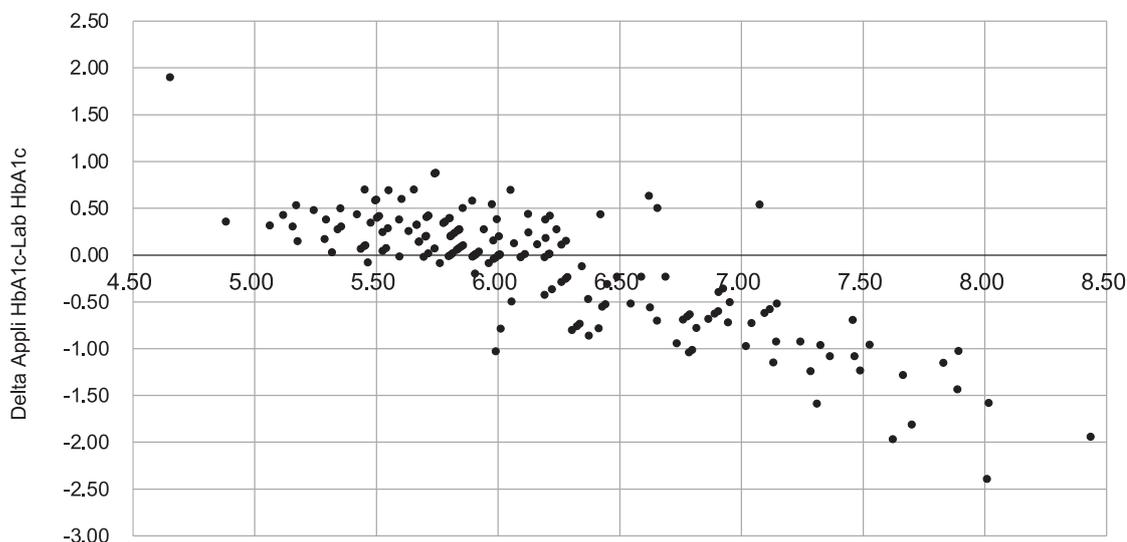


図 11

改良後アプリ HbA1c Bland-Altman 分析 (n=323)

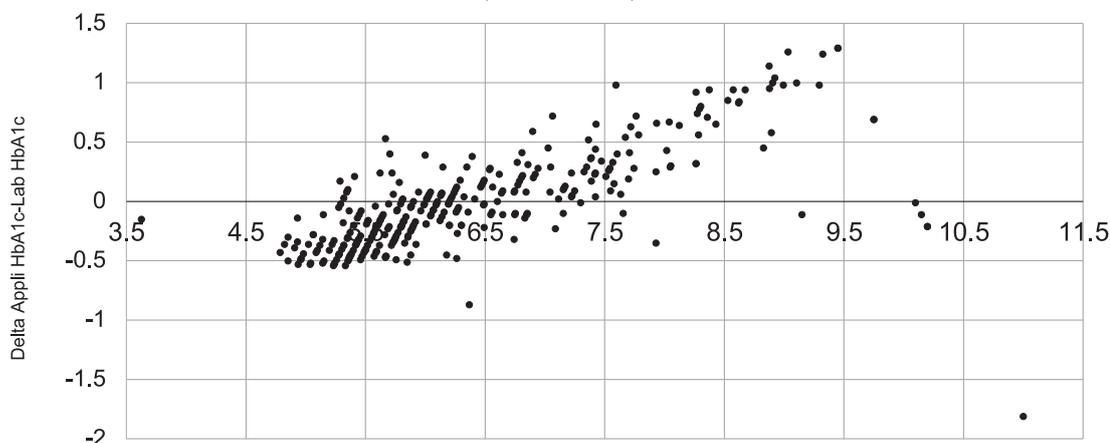


図 12

今後これが、たぶん奥先生のお話だと5年ぐらいで世界が広がってくれるといいなと思いますが、まだまだ十分なところではないということで、僕の研究をご紹介させていただきました。ありがとうございます。

大西 ありがとうございます。何かご質問がありましたらぜひ。

奥 顔認証の話はとても興味があって、私も最後までプロジェクトにならなかったのですが、ある企業が顔の実年齢とAIから診断した年齢の差が、要は若く見える人のほうが健康なのではないかという仮説があって、中国に論文があるんです。それを基にビジネス化しようと3年ぐらい前に検討したのですが、結局それは社内の優先順位で落ちてしまった。それ以来、私がもともと放射線科医というのがあって、顔の画像から何ができるのかということはすごく興味があって、今ヘモグロビンの話をとても興味深く聞いたところなんです。どの辺だったらできそうという、今のアウトカムとしてはご説明していただいたとおりでと思いますが、やればこの辺はできるのではないかと考えていらっしゃる場所はどの辺ですか。

下澤 皮膚と老化はすごく大きなテーマで、老化研究をやっている先生たちは結構いらっしゃるんですね。皮膚の細胞レベルではいろいろなことが分かってきているのですが、それは見た目ですれぐらい違うのかは、しわの数や色素沈着ということで、いくつかのパラメータを入れればできてくるのではないかと思いますし、わりと簡単にできるのではないかと思います。

一番の問題は、黒人や色素が違う人たちをどうするかということ인데、僕らは皮膚の表面を見ているのではなくて、血流を見ているので、黒人でも測定できます。じゃあ、顔ではなくて手のひらでもいいのではないかと問われたら、確かにおっしゃるとおりなんです。手のひらでもいいけそうですが、顔のほうが自撮りしやすいし、いいかなということで顔認証としていますが、手でもできるのではないかと考えています。実際、酸素は爪でやりますからね。あれは赤外線を使うので別ですが、これは普通の可視光でやっているということです。

柳原 これすごく面白いのですが、これってどんな方法で顔を撮ってA1cを？

下澤 それは特許の都合で。

柳原 言えないですね。やっぱりお話が出たようにやっぱり血流。A1c、どんな種類なのかなと思ってですね。

下澤 完全に血流なんです。

柳原 顔で血流ってなかなか、本当に表在の血管しか見えないので。

下澤 そう、表在の血管。毛細管の中の血流ですね。大血管だと駄目なんです。

柳原 眼底だったらもうちょっと精度が上がったりしますか。

下澤 でも眼底は見られないので。眼底を見るのは大変なので。散瞳して眼底でこんなことをやれば、もっと正確に見える可能性はあるかもしれません。

柳原 手間ですもんね。

下澤 手間です。それでは意味がないので。

柳原 確かに1でしたっけ、5プラスマイナス1ぐらいで出ましたけれど、血流でそこまで絞れるのかなど。

大西 すごいですね。本当にすごいなと思って。

下澤 彼らは工学部の先生たちだったので、20%の誤差はすごくないですかと言ったので、それは僕らの世界からしたら話にならないですね。

奥 おっしゃるように、SpO₂を測るものでも赤外線ですよ。結局先ほどもちらっと言いましたけれど、iPhoneで撮ることは可視光にいったん限定しているわけで、もしかしたらそこをもうちょっと緩めるとブレイクスルーがあるのかもしれない。

全く違う観点ですが、私は放射線科医なのでCTとかMRIなどは身近にあるのですが、今のセキュリティとか情報保護の問題で「次世代医療基盤法」という法律があって、国も治験のデータにビッグデータを使っていくということです。別の文脈ですが、倫理的に、例えば希少疾患で、プラセボに分類される人は不利を受けるので、プラセボ群をそもそもなくそうという議論が今ありまして、その究極はビッグデータから対照群を持つてくる。実薬群しか定義しない。そうすると不利益を受ける人がいなくなる。

そういう議論をしているのですが、今何を話しているかということ、そのときにCTやMRIで顔が写っていると、どんなに他の文字情報をマスクしても顔から全人類が特定できてしまう。そこを次世代医療

基盤法の画像を使うところの、今、内閣府や厚労省などいろいろ関係しているところでは、そこをどうやって個人識別をできなくするかという議論をしているのですが、これが相当難しい。

結局そこをなくしてしまえばいいのですが、なくしてしまうと、例えば皆さんよく見ているCTやMRIをじっくり輪切りにしたものでさえ、本当は集めて冠状断を作れば顔が再現されてしまい、その画像にどういう加工をすればいいかという複雑な議論になってきています。なので、顔を撮るということに関してはそういうリスクがある。ただ顔ではなくて、皮膚だったら、手のひらだったらいいのかというと、それでも本当はなくて、手のひらでも同じ議論になる。

下澤 指紋がありますからね。

奥 結局画像のところは、先ほどのストレージの話とも関係ありますが、あればあるだけそういう問題が起きてきて、21世紀の初頭に生きている人間としては、もうその話をやめてしまおうよと。個人情報一切やめてしまおうと、話がシンプルでいいのではないかと思っているのですが、そのために国民的なコンセンサスとか世界的なコンセンサスがあるかなということ、そんな問題も。

下澤 この場合、顔認証、今、全部の顔を撮っていただけますけれど、ここの血流だけなので、要するに虹彩が写らなければ個人の特特定は今のところは無理ですよ。

奥 今のところは難しい。

下澤 ここの皮膚の一部だけの写真であれば、おそらく個人の特特定はできない。手は駄目なんです。指紋もあるし、掌紋もあるので絶対使えない。でもチークだけだったら、今のところの技術だったら個人は守れるだろうと。僕らはなぜ最初に個人向けのアプリを売ることにしたかということ、個人は自分の責任で自分の顔の写真を撮りますから、それをSNSに上げたらそれは自分の責任なので、このアプリの責任にはならないということなんです。

奥 ご参考の一つだけご説明すると、そういう意味でビッグデータの事業者の中で、医療側でやっている人よりも、もうちょっと医療側でないところでやっている人たちは、情報を全部返してしまう。返却と呼んでいますが、返却していったん個人の所有にして、返すときに、だけど使わせてねと。一回返

して、すぐ戻してもらおうということをやると、結局個人識別の問題が全部クリアできてしまうので、そういう返却という概念をもう少し浸透させたらどうかという話があります。その辺はそれを言っている人たちとどうやっていこうかと。

技術としては、皆さん聞いたことがあると思いますが、ブロックチェーンとかNFLとか、そういうタグ付けする技術で、個人に返したものを全部もう一回戻してやるという。個人に帰属するデータのフルセットの集合として何かビジネスを。今の顔認証もそうですが、という議論がこの1年ぐらいで始まっています。

下澤 これは倫理申請の時にすごく問題になりました。顔写真を撮らせてもらうということで。

大西 先生がさきほど、診察室で音を録るとおっしゃいましたね。同じように顔の画像を撮ることもできるわけですよね。そのときに今のHbA1cのデータとの違いは果たして意味があるのかなと。糖尿病の人がこういう顔をしていたら、それは糖尿病。だからHbA1cのデータよりAIで撮った画像データのほうが実は正しいみたいな世界になってくるのではないかと。あくまでHbA1cは測れるから測っているだけであって、HbA1cがちよっと高いからといって、その後の合併症が大きく違う人がいっぱいいるじゃないですか。だからどっちが正しいのか、微妙かなと感じます。

奥 もうひとつ、画像、僕はもともと放射線科医になったのもその辺ですが、画像は今、中央放射線部で管理していますよね。あの分量と、波のデータ、例えば音声や心電図などと桁が全く違うんです。今の最先端のストレージ技術でも画像全部は持てないんです。まだ持てないんです。だからもちろん東大病院は、東大病院で撮った画像は持っていますけれど、それが例えば診察室のビデオになってしまうと、急に画像の情報量が増えてしまうので、たぶんそこは僕が知っている範囲でもあと10年とか20年かかりそうな。量子コンピュータとかいろいろな技術でそこはブレイクスルーが出てくればいいのですが、今の技術の延長上では画像が全部持てないということのようです。

大西 そのあたりが、さきほど先生にお伺いしたことにもつながってくるのですが、やはりエネルギーコストがネックになるのかなと。もちろん、技術の

進歩で将来的には大丈夫だと思うけれど、そこがネックになるのかなという不安はあります。

奥 私が放射線科医になった1988年には、CTを10枚くらいしか撮らなかったんですよ。今は何千枚と撮っている。でもそれが本当にその状態で情報が必要なのかというところと違うので。

大西 HbA1cだけでなく他にもいろいろな検査値が顔認証でわかるようになる可能性があるのがすごいと思いました。

下澤 ただ可視光はやっぱり限界があるんですよ。今、赤外線カメラがすごく安くなりましたよね。あれをうまく使えないかなと思うんですけど、スマホに赤外線カメラがまだ入っていないので。私は竹中工務店との共同研究もあるのですが、竹中工務店は何をやっているかというところ、AIを使って患者さんのベッドからの転落の予知をやっていて、それは赤外線カメラでデータをとっています。赤外線カメラなら真っ暗な寝室でも撮れます。

奥 人間は難しく、下世話な話ですが、可視光ではないものが撮れるようになると、例えば裸が撮れてしまうという話になる。そこはそこで技術的なブレイクスルーはうまくいっても、法的になかなか難しかったりする。特にスマホというコンテキストでいうと、可視光を超えるものを掲載することはかなり大きな議論になる。

柳原 空港のあれでも見えるんですよ。空港のこんなやつも結構見える。

奥 裸に近いものが見えてしまうので、あれは入国審査という特殊な状況で許されている。

あとですね、放射線科医としての発言としては、可視光と赤外線などを組み合わせれば全部裸が撮れてしまう。物理量AとBの差があれば、結構全部分かっちゃうんですよ。そういうものなので。禁断の技術です。

下澤 寝室の画像を撮ろうと思ったら、赤外線カメラが絶対必要なんですよ。上半身の寝返りの、こういう動きをすると落ちるといいます。それを覚えさせて、この人はこっち側に寝返りを打つなら、そっち側に壁を置いておけばいいだろうと。そうすると落ちなくなる。

奥 リミテーションがあるという発言をしながら、介護のDXという別の文脈で言うと、そここのところはものすごく重要で、そこがクリアできれば。要す

るに転落防止と褥瘡防止。その二つができることが介護DXの必須要件。

下澤 スマートハウス構想で、オムツの中にセンサーを入れておいてというのがあるんです。それは科研費Aから取っています。看護学科がやっています。

堤 一つだけいいですか。顔の写真で、HbA1cだけでなく、尿酸だったり、TGだったりも併せて検査できていますが、見ているものは違うのですか。同じ血管を見ているのでしょうか。

下澤 血流の流れの解析の方法が違うんです。スペクトラムが違う。さっきのLC/MSみたいなものです。ピークが違うんです。A1cのピークとコレステロールのピークが違う。

堤 難しいですね。

下澤 波形が、その中のどこのピークを見るかによって。イメージ的にはそういう感じです。

堤 色の黒い方とか、しわが多い方など問題ないのでしょうか。

下澤 今のところ90歳まで大丈夫です。

柳原 技術の進歩がすごいですね。

下澤 年齢による差はあまりないようです。

柳原 血液検査が要らなくなるんじゃないかと。

下澤 他のいろいろなものを全部測れるようになったらすごいと思いますけれど。白血球とかまで測れたらすごいですよ。

大西 下澤先生、ありがとうございます。最後は、東京大学大学院医学系研究科衛生学分野の石川俊平先生に病理領域におけるAIの現状、問題点および今後の展望についてお話をいただきます。よろしくお願いたします。

病理領域におけるAIの現状、問題点 および今後の展望について

石川 石川でございます。早速始めさせていただきます。

病理画像領域のAIの現状と、技術的なものもあり、いろいろ課題としてあるのですが、全体として放射線領域に比べてあまり進んでいないのが現状だと思います。たぶん、それは、例えば放射線や内視鏡はもともと画像を撮っていて、すでにデータがあるのですが、病理はスライドを1回作って、そこか



石川 俊平 先生

らまた撮影しないと駄目で手間が二重にかかる。そういう意味でデータを集める現場の負担も多いので、それなりに放射線や内視鏡に比べて進みが遅いのかと思います。

Singularity というか、どの世界にも人間を人工知能が超えたというのがあると思いますが、病理画像については 2017 年の JAMA の論文で、リンパ節の中の乳癌の微小転移を領域で見せているものがありますが、それが人工知能がこういう微小転移を検出するという、そのコンペでその当時の State-of-the-Art の人工知能のプログラムが普通に人を超えてくる。病理医も 30 時間かけていいよという、だいたい同じぐらいになってくるのですが、こういうのは、人は普通に疲れるし、ミスも起こるので、こういうコンペをやると AI のほうが必ず勝ってしまう。ただ、これが 2017 年ですが、これが人を超えたという報告が正式にいうと最初のスタートなのかと思います。

今、いろいろなプログラム医療機器が病理の世界でもあるのですが、放射線や内視鏡領域のほうが明らかに多くて、少なくとも日本ではまだないんです。厚労省としてプログラム医療機器として認められたものはなくて、いわゆるデジタルスライドによる診断というのが認められているぐらいになります。

ただ海外にはいろいろなプロダクトがあって、代表的なものが Paige と言いますが、前立腺に小さな病変がある場合、今は一人に十何本、生検において探すのですが、この Paige という企業がいろいろな比較試験をして検出率が上がると。ただ、こういうのをやると偽陽性が往々にして上がるのですが、偽

陽性の上昇がなく検出率が上がったということで、今プロダクトとして認められています。ただ、いわゆるセカンドルックとって、これだけに任せると危ないということで、一応病理医が全てのスライドを見て、パラレルに走らせて見逃しを防ぐという意味なのかなと思います。

これに類するものが、大腸癌の MSI のカテゴリーがあるのですが、そういう診断をするものがいくつかありますが、プログラム医療機器として認められているものはあまりないです。

実際、いろいろな病理の先生のユーザーエクスペリエンスを聞くと、画像を撮らなければならないこと自体が、つまり普段顕微鏡で診断しているので、それをわざわざホールスライドイメージを持ってきて、それをプログラム医療機器の中に移さなければ駄目というのは非常に手間がかかるので、そういうのが疾患ごとにいっぱいあるわけです。前立腺がんの例になりますと、いろいろな癌を普段診断しているので、それを全部、ユーザーエクスペリエンスを保ちながらやるというのは結構難しく、なかなか浸透しないのかなと思います。

こういう人工知能はいろいろありますが、病理の人工知能の独特の難しさは、技術的な話になりますが、他と違うのは、一般画像だと例えば家とか猫を判定するのにだいたい小さなサイズの画像があれば十分だと思いますが、病理というのはホールスライドイメージという大きな画像があって拡大すると、細胞一個一個が見える画像が必要になります。全体を見つつ、細胞一個一個を見なければ駄目というのは、本質的にもものすごく大きなサイズの画像を扱うことが必要になるので、普通の一般画像とはちょっと違うことになります。例えば猫を判定するのに毛を一本一本見る必要はないと思いますが、病理の場合だったら必要になりますし、そういう大きな画像データがデフォルトで、医療現場で集められていないというのがあって、それもあってなかなか進みが遅いというのがあります。

ただ、いろいろなアプリケーションが出てきて、例えば複雑な形態情報をニューラルネットワークの中にエンコードして、癌であるとか癌でないとか、どういう癌であるかを見るというのはありましたが、もっと一般的に病理画像を数値として単純に表現することでいろいろなことができるようになって

きました。難しい表現ですが unsupervised アプローチといって病理画像を単純ベクトル表現すると、普段こういう病理画像は専門家の人しか扱えないのですが、普通の検査値のような感じで数値データとして扱っていただく(図13)。例えばサブクラスを同定したり、類似症例検索をしたりと、他のデータと合わせるように、何かこういう数値表現を人工知能でやることによって、いろいろなアプリケーションに使うことが最近出てきています。

8,000 例ぐらいの癌の病理組織を二次元にデータを圧縮して数値表現したものがあありますが、一つの小さい四角が一つの症例に対応し、皆さん普段見ていらっしやらないかもしれませんが、確かに拡大すると、数値が近いと、より組織像が似ているのがなんとなく分かると思います。例えば高分化の扁平上皮癌が固まっているとか、高分化の管状腺癌が固まっているというのがあります。

1万症例近くの症例のダイバーシティというのを見ると、なかなかこういうのが人の脳ではできなかったんですが、こういう病理組織像という形態を数値化することによって、可能になります。

これができると何ができるかという、例えば同じ腺癌でも、形の分布をこうやって見せると、典型的な青色のものがある中で、ちょっとその外側に、

例えば何か変わった形のものがありますと。例えば Neuroendocrine のサブタイプがあったり、オレンジ色だと大腸型の腺癌みたいなものがあるのですが、要は何か変わったサブグループがあるということが病理の知識なくして、事前の知識なくしてこのように分かるということがある。

あとは胃癌のバイオプシーで、よく似たグループ、免疫療法にレスポンスの異なるグループの症例が分離されてきたり単純にその形態を数値化することで普通のデータとして扱うことが可能になってきています。

その中で期待されているのが類似画像検索というものです。例えばさっきお話したように、数値が近いものをいろいろなデータベースの中から取ってくると、こうやってクエリ画像として、何か変わった症例があったときにデータベースからそっくりな、おそらくそっくりだなと思っていただけだと思いますが、そっくりな画像が今まで別の症例であったり、他の病院であったりということが分かります。

普通こういう検索をすると、例えば形を表すと、人の顔だったら似た顔の人を探してくるのに、長い鼻や大きな目など一回言葉にして検索するのですが、そういう表現するとなかなかうまくいきません。ところが形の特徴で直接検索するとすごく似た症例

AIによるがん病理組織像のユニバーサルな数値化

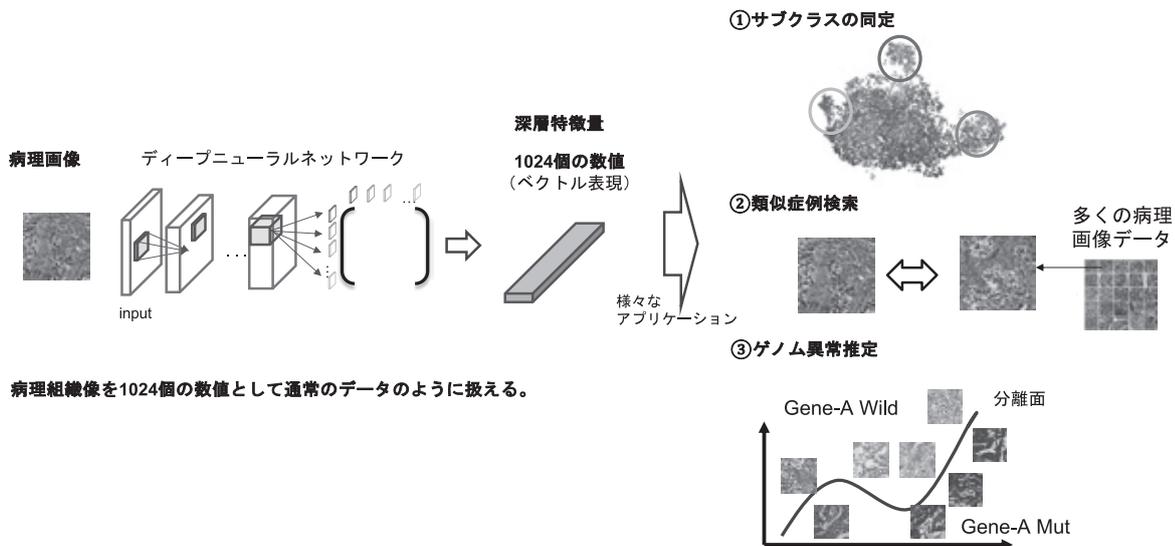


図13

(図13は巻末カラーで掲載しています)

があって、過去にどういう症例だったかが分かる。あとは不明な症例があれば診断名も分かります。

今、WHO 分類の中に癌の組織カテゴリーは自分で重複を除くと 1,100 種類あるらしいですが、高尾山の山の植物の数よりやや少ないぐらいということです。山に入って植物の名前が全部言えるぐらいの、それぐらい複雑な作業なんです。経験のある病理医も分からない画像は必ず出てくるんです。そのときに、周りの方に聞いたり、アトラスをめくって探したりというのをやるのですが、そういうときに類似画像検索が使えらると。今、いろいろな類似画像検索のアプリケーションのために人工知能が開発されています。

実際に類似画像検索をしたものと、していないもので診断精度を比べると、シニアのクラス、ジュニアのクラス、レジデントのクラスなどいろいろなクラスで効果があって、全く分からない人でも普通の病理医ぐらいになることが分かります。

途上国での話になりますが、途上国ではこういうデジタルデバイスがないんです。今はスマホに実装して途上国で使ってもらっています。病理のスライドスキャナーが 1,000 万円以上するので、日本の中でもこれを全施設に入れるというのはリアリティのある話でもなくなってきている。こういう一般的なデバイスに、特に医療のリソースのないところで使うことが。私のグループでもそういうのをやっていますけれど、いろいろなところでこういうことが試みられたりしています。

例えばラオスでは、いろいろな染色液がペットボトルのようなものに保存されており、病理の染色態度というのは病院によってまちまちなんです。たぶん日本の中でも病院ごとに違いますし、東日本と西

日本で売れ筋の染色液は違うんです。

人が見るとカラー写真でも白黒でも同じような情報を取りますが、AI にとっては色のバリエーションは、ドメインの差といいますが、かなりノイズになるので、こういうことをどの施設でも同じような結果を出すことが、おそらく放射線画像以上に難しいのではないかと想像しています。どんな色でも人間に診断できたので、あまりそこをクオリティコントロールしようというモチベーションはなかったのですが、こういう人工知能の研究が進んできて、こういうのがかなり難しいことだということが分かるようになってきました。

例えば、同じ画像を別のスライドスキャナやスマホで撮ったものについて、これだけ違います (図 14)。染色法によっても、この中に 32 万枚ありますが、同じようなものを撮ってもいろいろなパターンのものがあります。おそらく臨床実装するのに一番大きなハードルになっているものの一つが、施設間差というものです。同じようなパフォーマンスをどの施設でも出せるか。クオリティコントロールが結構難しいと思っています。

もう一つは、今言っているような、一般的に数値にするというのは深層特徴量と言われていて、非常に深い表現力、例えば癌らしさとか炎症のシビアさなど、人の顔でいうと男らしいとか女らしいとか、ものすごくふわっとした表現なんですけど、これは何を AI が認識しているか分からないという、いわゆるブラックボックス問題があるので、もう少し説明可能な特徴量を出すことも行われています。

それが今われわれも開発しており、HE 画像からどのような細胞種が、これだけありますみたいなことです。生物学的なエレメントにきちんと推定する

デジタル画像機器による同一病理組織の画質の違い

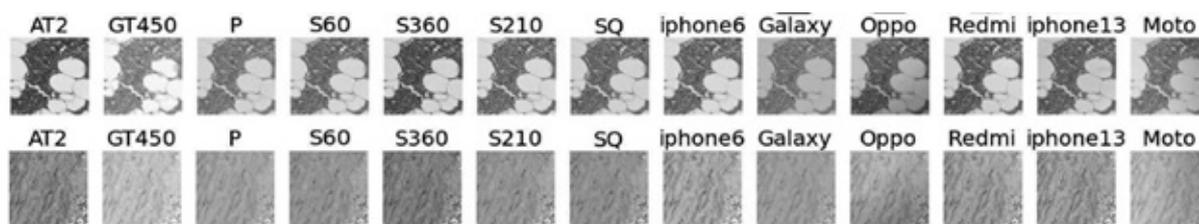


図 14

(図 14 は巻末カラーで掲載しています)

ようなAIがあって、ここでは例えばHE画像と剥がして免染するみたいな、こういうペアのトレーニングデータを積むことでこういうことができます。こういうことをやって、全ゲノム解析等実行計画のような中でわれわれもいろいろな医療機関から病理画像を集めて、解析結果を説明可能な状態にして返している。例えば腫瘍の割合はこれぐらいであるとか、リンパ球はこれぐらい浸潤しているみたいなところを返して、今は研究レベルですが、こういうのもエキスパートパネルで議論できるようなレベルのものもっていきようとしています。

今、お話ししたことを、空間ゲノミクスみたいな話も持ってきていますが、今日は省きます。

今日お話ししたところをもう一度確認していきますと、深層特徴量を利用して、最初にAIの現状をお話しした2017年にAIの機能が病理を超えてきたというのがあって、いくつかのプロダクトがありますが、あまり放射線領域や内視鏡領域に比べて進んでいないというのがあります。

癌か癌でないみたいなアウトプットのすっきりしたもの以外にも、一般的に病理組織像を数値化することで、普通のデータサイエンスのように組織学にいろいろなアプリケーションをつくるのが今出てきて、今日お見せしたのはサブタイピングとか類似画像検索、あとは病理画像から体細胞変異を推定するのもよくあるアプリケーションです。それを一般的にはスマホのようなところに載せて、医療資源の乏しい地域で使おうというアクティビティもあります。それ以外に深層特徴量以外に、こういう細胞種のような非常に分かりやすいエレメントを作るといって、こういうAIもあって、使われようとしていることを簡単にご紹介します。今日は十分に練られていない資料で申し訳ありませんでした。

大西 ありがとうございます。ご質問はございますでしょうか。

奥 一つ感想と質問です。感想のほうは、先ほど8,000例の写真を1枚、2次元に落とすというもの。これはとても僕の中でも興味深くて、これが例えば各国の保険施策に使えとか、日本だけでも癌のどの分野に重点的にお金を出すとか、社会医学的な、公衆衛生学的なアプローチに使えのかなというのが感想です。

質問は、ブラックボックスの問題を先ほどだいたい

議論したのですが、先生の中では説明性はどのぐらい重要だと思っていらっしゃるのかということ、そこをできるだけフラットな形で、説明がそもそもの程度必要なのか、必要ではないのか。必要だということなのかと思って聞いていたのですが、そこを教えてくださいませんか。

石川 確かに今おっしゃるように、ブラックボックスという議論があって、それをなるべく可視化しようという技術は出てきている。一つそれをまずお話ししたいと思いますが、例えば犬と猫があって、犬ということにすごい寄与する、AIが認識する領域はどこですかみたいな、犬の顔が判断材料になりますね。猫だと顔だけではなくて、しなっとした体や模様が重要だということは分かるのですが、TP53の変異を判定するのに何を認識しているのか見るのに、例えば核の異型みたいに核に染色体異常が起こってゆがんでいるとか、免疫原性の強い胃癌だと上皮の中に入り込んでいるリンパ球を認識しているというのが分かります。

こういう技術は出てきていますが、これを見たときにリンパ球の浸潤だけで説明できるかということ、AIとしての切れ味というか情報量としては落ちてしまうんです。説明可能なものだけで何か全部をやらせると情報は落ちてしまっていて、どうしても深層特徴量というふわっとしたものの方が性能がいいんですよね。

ただ、説明可能性が大事だという技術的な理由としては、いろいろなAIはどうしても局所最適化とか、アルゴリズムでいうと数学的に、全然生物学と関係ないところを認識して変な判定をしてしまう。病理画像だったら、例えばこういうことも考えられて、癌と正常ですごく識別性能がいいのですが、実はラベルのところの駒込病院であることを認識して癌の検体を集めていた病院だと、そういうのを見ている場合があるので臨床で使うと思ってもよらないアクションを起こしたりするんです。そのときにちゃんと認識していますか、ということをやるといのが、おそらく規制のような観点では必要なかなと思うんです。

ただ、結局さっきみたいに、こういう細胞がこれだけいますからこうですというほうが、より生物学とか、いろいろな研究としては広がりがあるので、説明可能性というのは非常に大事だと思います。そ

ういう二つの意味で大事なのかなと思います。生物学的には広がりという意味と、規制、いわゆる局所最適化をちゃんとコントロールする意味では説明可能性は大事なのかなと思います。

奥 ありがとうございます。

大西 素人質問で恐縮ですが、そもそも私は画像をAIが認識するときは、全部数値にして認識するのかと思っていたのですが、先ほどおっしゃった1,024個のベクトルに分けるといのはちょっと意味合いが違う感じでしょうか。

石川 今日ご説明した病理組織像を、1,024個のベクトル表現のようにするというのは、もう少しどういう意味か説明すると、例えばこういうAIがあって、病理画像を見せて、最初、数値はRGBの単なる色の情報ですが、どんどん深層ニューラルネットワークの深いところに行く最終的には癌らしいとか、正常らしいみたいな意味になるのですが、その途中みみたいな、中間の数値を取り出しているというように考えていただけたらいいかと思います。

ひとまず、病理組織像を数値として表現しましょうというステップがあって、特徴抽出とありますが、その後判定という二段構造になっているので、まず特徴を抽出して、その後いろいろなアプリケーションを使うのが今回のこちらの手法なのかなと思います。

これを使って、その後癌か正常かという判定もできますし、もっと数値としてそのまま使って、サブクラスを同定したり、類似検索もできるのかなという理解ではいます。

奥 少しだけ補足しますと、画像を石川先生がおっしゃったように、数値に落とすということを最近はしない領域の研究が増えています。画像は画像のまま。というのが情報科学の中では一つの大きな分野なので、こちらのほうが人間には分かりやすいんです。数値を経て、ダウングレードする必要もないような。さっきの先生のご説明で猫のしなやかさとか、どこを見ているか分からないので、ある意味これは1,024個の決めつけなので、そこはしないという研究は結構最近は多いです。

石川 今おっしゃっていただいた感じで、ここからの変換の仕方はいろいろな研究があって、トランスフォーマーとかいろいろな名前が使われていますけれども、それぞれ特徴があります。今日は局所の細

胞、これでも1,000個以上あるのですが、ホールスライドの中のごく一部の局所だけで判定するのですが、実際病理の情報はもっとスライド全体の、この辺で病気ができて、このように進んできたとか、いろいろな情報を読み取るので、もっと広い領域を読むようなアルゴリズムがいろいろ出てきて、それらを組み合わせでいろいろなことができるようになってきているかと思っています。

大西 ありがとうございます。

堤 今のお話を聞いて、病理領域はまだAIにいろいろなハードルがあるため、まだそれほど進んでいない。例えば染色法とか染色液による染色の程度が一つネックになっていると理解したのですが、何をスタンダードに置くか分からないですが、AIがそれを補正して標準化するとか、何かそういうことはAIにできないでしょうか。

石川 まさにそのとおりですね。例えばこのように、同じものでも撮影によってこれだけ違うというものもありますし、単純には色のノーマライズを使ったりします。ところが実際やってみると、色って人間には分かるので色が違うなって感じで、実際人間が簡単に分からない高周波の成分とか、何かの特徴が撮影機で違ったりするので、色だけを変換してもなかなかうまくいかないんです。例えばホールスライドをスマホ風に変換するとか、色を含めた画質変換みたいなことです。そういう意味で、今おっしゃっていただいたことは広い意味で正しいのですが、変換してなるべく近づけてやるようにはしています。

ただ限界があって、プラクティカルにはAIにいろいろな施設とか撮影機の画像をきちっと読み込ませる必要があるかと思っています。逆にそれができるれば、だんだんこういう施設間の差はなくなりますが、最終的にはそれぞれの施設でファインチューニングというんですか、少しパラメータを変えて、LDTみたいな使い方をするのがプラクティカルなのかなと思っています。

堤 まだいまひとつというところで、そこがクリアされると格段に進歩するのか、あるいはまだ他にもハードルがあるのでしょうか。

石川 AIモデルという意味では、いろいろな画質の影響をクリアすれば、ある程度完成度が高いものができるのかなと思います。実際にそれを使うこと

になると、いろいろなユーザーエクスペリエンスの問題、病理は普段診療していて、たぶんいろいろなものが選択肢に入るので、それぞれのAIのアルゴリズムのようなものを個別に動かしたり、そこに画像をもってきたりするのが実際は難しいので、そこは統合的にユーザーエクスペリエンスをものすごく良くする必要があるなと感じています。

あと、正直いろいろな内視鏡や放射線に比べて、ユーザーの数という意味でも、なかなか企業側の開発モチベーションがわからない領域であるのは確かだと思うんです。そういう意味で、普通の病理医のまねをするようなアルゴリズムは、コマーシャルとしてもうまくいっている例がありません。

最近その代わりに出てきているのは、もっとハイエンドな、例えばいろいろな新薬の治験などで、普段病理医が見ないような、それこそ免疫療法の効果判定とか、いろいろな新薬、バイオロジックの効果予測したりというものです。あとは例えば、治験をやる時に、年齢や性別を一致させますけれど、病理組織像で背景を一致させて、もうちょっと各群のバリエーションを少なくして統計的な有意差を出してくるとかです。そういうときに同じ染色法で同じスキャナーを使って、もっとコントロールされた条件でAIを使うとか、そっちの方向はわりと広がりが出てきているのかなという印象は持っています。また病理はマンパワーが足りない世界なので、AIで作るという一時的なハードルが、普段の診療の中で乗り越えられない領域なのかもしれないです。

堤 ありがとうございます。

下澤 先生が出されたのは基本 Histology（組織診）ですよね。Cytology（細胞診）と比べるとどれぐらい変わってくるものでしょうか。

石川 Cytologyのほうが難しいと思います。でも需要はかなり大きいと思います。Cytologyはスクリーニングで多数の症例を処理しますし、あれを今は人が見ているので、その需要はかなり大きいですが、Histologyと違うのは、たぶんご覧になられたことがあると思いますが、焦点が合わないというか、細胞がぐちゃっとなって非常に形態が見にくいんです。スミアでやる場合はそのようになるので、最近は液状化細胞診みたいなものできれに行う方法がありますが、世界のほとんどは、まだ焦点を検査技師が合わせながらやっています。焦点が変わると

AIはものすごく違う認識をするんです。画像も Histologyだと一枚になりますが、細胞診だと5つぐらいZスタックを行うので、データ量としても検索しないといけない領域も明らかに多い。ただ需要が非常に大きいところなので、いろいろな開発がされています。いろいろ課題は多いと思いますが、データ量の課題は大きいと思います。

下澤 僕は尿沈渣を、Cytologyなので、数値化までもいかないし、とても難しいなと思っています。

石川 プラクティカルには、絶対にここには異型細胞がありません、というところをルールアウトしてくれるだけでも助かるかもしれないです。あとはもう検査技師がやるという。全部を今スクリーニングしている状態なので、何か見つけるというよりも、絶対違う領域をルールアウトするようなプログラムがプラクティカルなのかなと思います。

下澤 あとISO的にいうと、学問的でなくて申し訳ないですが、検査室の役割というのはもちろん診断を下して、それをクライアントに返すことも重要ですが、クライアントに次に何をすべきかというサジェスションも重要。例えば、こういう結果があるから免疫染色、こういうのをしたほうがいいのではないとか、こういう検査を追加してみたいかがでしょうかみたいなことをするというのもわれわれの仕事のうちの一つになるのですが、そういうところを病理のAIから次のサジェスションはできるものでしょうか。そういう研究は進んでいるのでしょうか。

石川 これはかなり正確なポイントだと思いますが、例えばどういうことに利用できるかという議論があったかという、病理検査室で、初回 HE 染色を作成して、その後いろいろな診断の選択肢があるので免疫染色をたくさん加えますよね。そのときにもう一回切片を切り直したり、作り直したりする。そこに TAT (Turn Around Time) として数日間、1週間かかるので、事前に診断ではないですが、免疫染色がたくさん出そうな組織かどうか、どういのが出そうかをある程度 HE 染色をしておいて、事前に準備しておく。そうすると、次のサジェスションのところまでいってはいませんが、プラクティカルにそういう TAT が早くなるというか、そういう議論は今まであって、非常にリアリスティックな使い方だと思います。

柳原 何切片、何個から見れるとか決まっているんですか。私も血球とか微生物で同じなので1個でいいのですが、病理だと1切片に10か所とか何か所見るんですか。

石川 病理だと全部だと思います。全部見て、どこに癌があるのか分からないので、だから画像が巨大になって、そこがちょっと難しい。

柳原 確かに見逃すわけにはいかないからですね。だから大きな画像になると。分かりました。

大西 まだいろいろ議論は尽きないと思いますが、ちょうどいい時間になりました。

今日は知的好奇心が非常に刺激されるようなお話ばかりで本当に勉強になりました。ぜひこれは魅力ある誌面にまとめたいと思います。ありがとうございました。

堤 本日は検査・診断領域におけるAIの現状と未来というテーマで、4名の専門家の先生方からお話をいただきました。奥先生からはAIと医療ということで、これまでのAIの歴史と、今後AIが医療においてどのように活用されていくのかをお話いただきました。中でもAIが臨死期に携わることについては色々な意味で考えさせられる内容だと感じました。柳原先生からは微生物検査領域におけるAIの現状と未来について、AIによるグラム染色診断や質量分析、インフルエンザ診断用のnodocaについてお話をいただきました。感染症診療だけでなく

AMR対策にもAIが活用されていくことは非常に重要であると思いました。下澤先生からは検体検査領域におけるAIの現状と未来について、AIを用いた希少疾患診断の尿検査や、採血をしない血液検査など、先生がご研究されている内容についてお話をいただきました。顔認証により尿酸やHbA1cの値が測定できるというのは本当に驚きで、実現すれば患者さんの身体的負担の軽減だけでなくさまざまな観点で医療資源の節減に大きく貢献できると思われました。最後に石川先生からは病理領域におけるAIの現状、問題点および今後の展望について、AIを用いたがんの病理組織学的診断とその限界などについてお話しいただきました。病理領域ではAI活用はまだこれから、という状況と理解しましたが、今後さらに発展することが期待されるようにお見受けしました。医療ならびに検査領域で現在AIがどのように活用されているか、そして今後どのように進んでいくのかを知ることができる大変有意義な座談会だったと思います。先生方、どうもありがとうございました。

文 献

- 1) 厚生労働省、資料2「保健医療分野におけるAI開発の方向性について」(2018年7月23日)
<https://www.mhlw.go.jp/content/10601000/000337597.pdf>



庭のホテル東京にて



微生物自動染色分析装置Mycrium® のワークフロー 対応材料種：血液・尿 ※その他材料も現在開発中

簡便な操作と自動染色からAI判定サポートにより夜間休日のグラム染色結果報告が可能になります

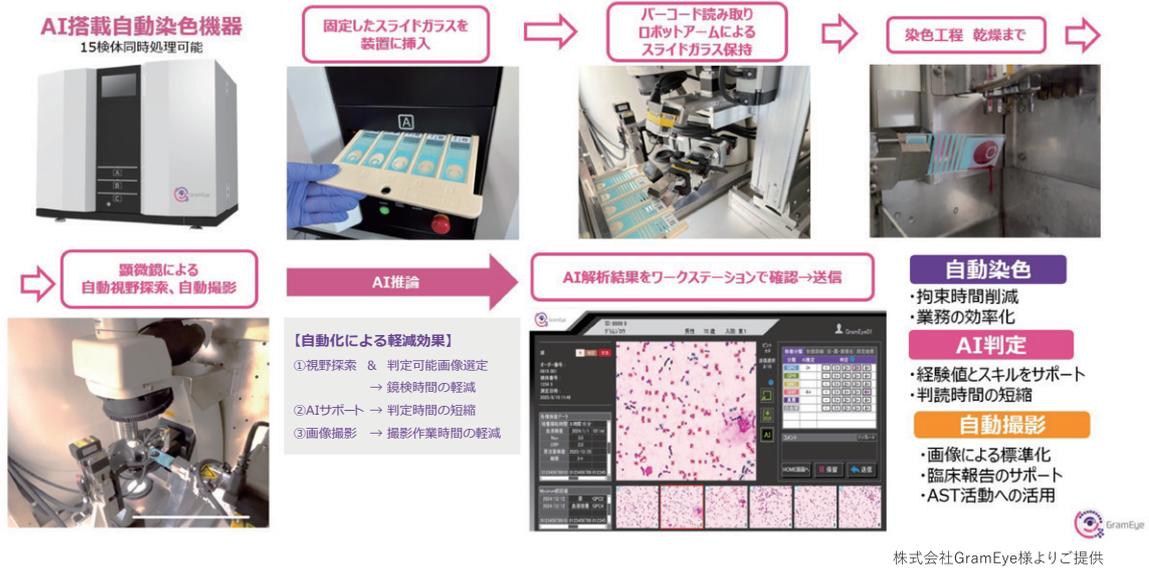


図 8

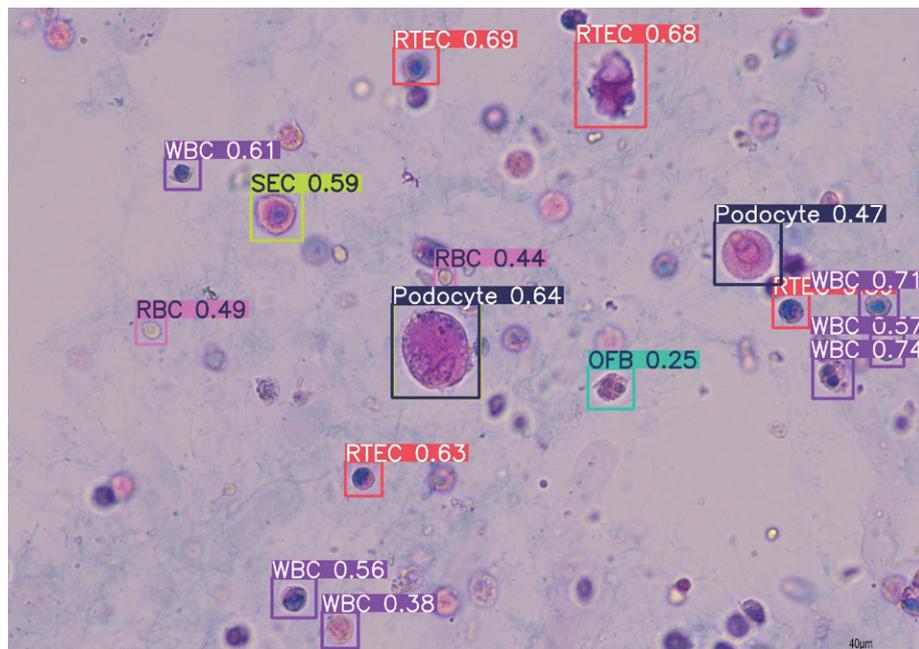


図 10

AIによるがん病理組織像のユニバーサルな数値化

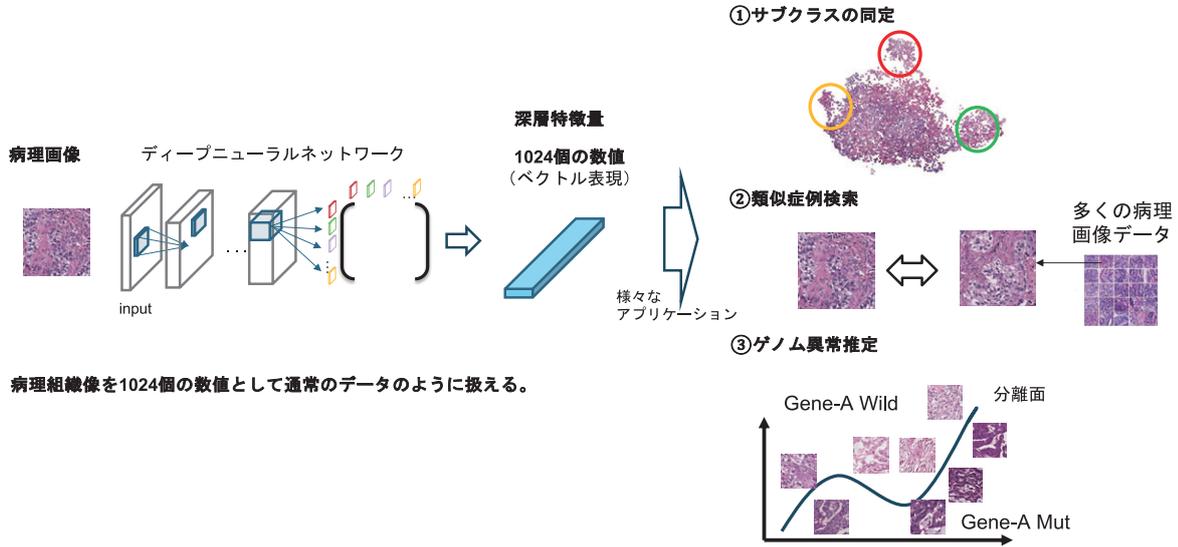


図 13

13

デジタル画像機器による同一病理組織の画質の違い

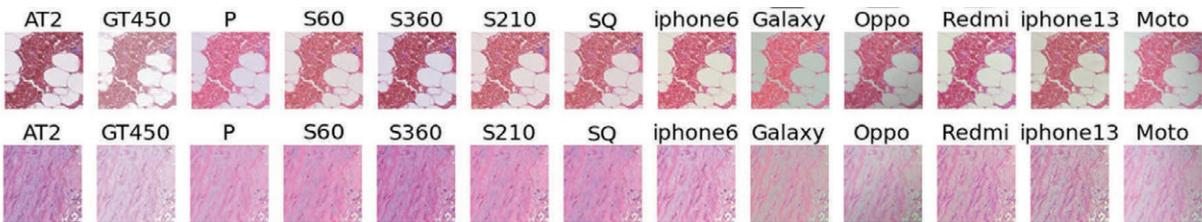


図 14