



Title	保健医療分野にAIをどのように活用することができるか？ : 糖尿病網膜症のスクリーニングを事例として
Author(s)	川崎, 良
Citation	makoto. 2025, 211, p. 2-7
Version Type	VoR
URL	<a href="https://doi.org/10.18910/101983">https://doi.org/10.18910/101983</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

# 保健医療分野にAIをどのように活用することができるか？

— 糖尿病網膜症のスクリーニングを事例として —

大阪大学大学院医学系研究科社会医学講座（公衆衛生学）・教授  
大阪大学医学部附属病院AI医療センター・副センター長

川崎 良

## はじめに

近年、見える形で、あるいは見えない形で人工知能（artificial intelligence, AI）が私たちの生活の様々な場面に活用され、広く浸透しつつある。それは保健医療分野においても同様であり、情報がデジタル化され、また、AIが数字から文字、そして画像、音声、動画といった様々な情報を扱うことができるようになり、その応用範囲が大きく広がっている。公衆衛生分野の保健医療の課題にAIを活用することに期待が集まっている背景には、医療現場が抱える様々な課題の解決のためにAIの応用が大きく貢献するという期待がある。すなわち、超高齢社会を迎えた我が国の医療においては専門医の偏在、診断精度の向上、そして深刻な人材不足の解消、また、医療だけでなく、保健分野においてもスクリーニングや検査判別支援、また、各種文書やコミュニケーション支援としての活用が挙げられ、これらの課題に対し、AIは大きな可能性を秘めている。その一方、その急速な進歩と応用範囲の広がりから未知なるものへの危惧がある。それは、人間の知性に迫る、あるいはそれを凌駕するのではないかという危機感さえ感じるような急激かつ高度な進歩が一気に加速しているからであろう。

本稿では保健医療分野でのAI技術がどのように開発されてきたのかを振り返り、現在、そして近い将来どのように活用されるのかについて、AI技術の活用が先行して進みつつあ

る糖尿病網膜症を事例に考えてみたい。

## 1. 保健医療AI黎明期（1950～1990年代） — ルールに基づくAIの誕生

“I propose to consider the question, ‘Can machines think?’”

*A. M. Turing (1950)*

数学者アラン・M・チューリングは、1950年の論文「コンピューティング機械とインテリジェンス」（文献1）において、「私は、『機械は思考できるか』という問題の検討を提案する。」という主張を展開し、人工知能研究の先駆けとなる概念を提示し、『機械』が人間らしいと言えるかを判定するためのチューリング・テストとして、相手が見えない状態で対話形式の会話を行い、機械と人間とが区別できないと審査員が判断する割合で機械が人間と同程度に知的に振る舞えるかを判断する試験を提案した。1955年には計算機科学者のJohn McCarthy（1927-2011）がダートマス会議（Dartmouth Conference, 1956）において、初めて「人工知能」という用語を使ったとされている。1966年には米国マサチューセッツ工科大学人工知能研究所のJoseph Weizenbaum（1923-2008）により、初の対話型自然言語処理プログラム「ELIZA（イライザ）」が開発された。ELIZAは、人間の入力文に対する簡単なパターンマッチによって応答を返すシステムで、臨床心理学者Carl Ransom Rogers（1902-1987）の来談者

中心療法 (Person-Centered Therapy) を現在で言うところのチャットボットで提供した先駆けとなっている。ワイゼンバウムは後に、「(ELISAが目指したのは、) コンピュータに最も適したタスクをコンピュータに委ね、人間の判断を要するタスクを人間に任せることである。」と語っている。Jones CR & Bergen BK (文献2) はELISAから最新の大規模言語モデルGPT-4.5-PERSONAまでのAIモデルのチューリング・テストを実施し、機械と人間とが区別できないと審査員が判断する割合がELISAは23%、GPT-4.5-PERSONAは73%であり、GPT-4.5-PERSONAは大学生被験者の69%を超える結果であったと報告している。

初期のAIの医学への応用の例として知られるのはMYCINである。MYCINは1972年にスタンフォード大学の研究チームが感染症の診断を行うことを目的に開発したシステムとされている。7年の年月をかけて集積された臨床の知識をデータベース化し、「もし、この症状Xがあれば、Y菌が起病菌である可能性はZ%である」というようなif-then形式のルールによる推論を返すものであった。臨床経験のデータベースに基づき、500個に及ぶif-thenルールを組み合わせ、症状から感染菌の起病菌を絞り込む推論は実際に当時の医師水準の診断精度をしのぐ性能で、「専門医の診断」の正解率(約80%)には及ばないものの、「細菌感染の専門家でない医師」よりは優秀な成績であった。このようなエキスパートシステムへの期待は高く、1980年代にはAIブームが興り、国家プロジェクトとして多額の予算が医学をはじめとする専門的な知識をもとに開発された。当時、大いに期待されたエキスパートシステムでしたが、大きな欠点があった。まずはエキスパートシステムの開発と更新に膨大な手間と時間がかかることであり、専門家の知見をルールとして収集し、検証し、手動でデータベース化していく

必要があった。2つめの欠点は、人間が使う曖昧な表現をエキスパートシステムではルール化できない点があり、例えば専門家の知見の中にはルール間で矛盾が生じるなどの課題が明らかになり、複雑な実問題に対して満足な結果が得られないケースがあった。また、臨床の現場での利用方法が容易ではなかったこと(MYCINの問いに対して30分かけて情報を打ち込む必要があった)やコストとえらえるメリットのバランスといった理由もあった。これらの欠点が次第にエキスパートシステムの発展に歯止めをかけてしまい、当時のAIブームは終焉を迎えた。

## 2. 保健医療AI萌芽期(2000~2010年代前半) — 機械学習の台頭

2000年以降には機械学習が保健医療分野に適用されてきた。特に、複数の要因、危険因子を背景として発生するがんや循環器疾患といった多因子疾患の発症リスクを評価する際に用いられる統計モデル、多変量解析を用いた判別や予測において、「訓練データ」をもとに機械学習モデルを作成するものである。すなわち、「過去のデータを用いて、次のデータを予測・判別する」モデルで、このことから「学習」という言葉が使われるようになった。決定木学習、サポートベクターマシン、クラスタリング、ニューラルネットワーク、などが代表的であり、現在でも広く用いられる。機械学習は、大量のデータからAI自身がパターンを学習し、予測や判断を行う技術で、ルールに基づくAIとは異なり、データに基づいてAIが自律的に学習するため、より柔軟で高度な処理が可能となった。その意味で、データに基づくモデル、データドリブンなモデルという考え方が生まれ、仮説に基づくモデルよりも、無仮説で大量のデータを分析し、複雑な関係性やパターンを認識することの可能性が示された。とはいえ、モデル作成にあたっては、どのようなデータをどのように利用す

るのかなど人間の設計が大きな影響を与える面もある。機械学習の登場により、画像診断へのAI応用が発展し、特に胸部レントゲン写真やマンモグラフィーの解析にコンピュータ支援診断（computer-aided diagnosis, CAD）システムとして応用された。機械学習によって数値情報や文字情報だけでなく画像の解析支援や判別・予測支援など応用の幅が大きく広がった。

### 3. 保健医療AI成熟期（2010年代後半～） — 深層学習革命

2010年代後半に入り、深層学習と呼ばれるAI技術が注目を集めた。ディープラーニングは、機械学習の一種であり、ニューラルネットワークを多層にすることで複雑なパターンを自動的に学習することができる画期的なAIモデルである。2024年にはニューラルネットワークによる連想記憶を考案したJohn Joseph Hopfield（1933生）とニューラルネットワークを多層化した深層学習を考案したGeoffrey Everest Hinton（1947生）がノーベル物理学賞を受賞したことは記憶に新しい。深層学習によって、画像認識や自然言語処理の分野で、これまでの機械学習を凌駕する高い精度の判別・予測が実現された。ディープラーニングの登場は、医療AIの可能性を大きく広げ、皮膚がんの判別や放射線画像の解析や診断支援、内視鏡画像・動画の病変判別支援、病理検査や微生物検査、また、創薬候補物質の探索、タンパク質の構造予測など、今や保健医療分野にAIは広く浸透している。自然言語処理においては、文脈に応じて柔軟に、そして、並列処理により高速に大規模なデータセットを効率的に学習できるトランスフォーマーモデルが提案され、その応用として現在広く使われているChatGPTの性能が広く知られるところとなった。会話や問診、また、カルテや医療文書、また、学術論文などを構造化してデータベース化したり、人間が

書くような自然な文章を生成したり、質問に答えたりすることができるようになり、保健医療現場での応用が多数報告されるに至っている。

### 4. 保健医療分野でのAI応用の事例：糖尿病網膜症

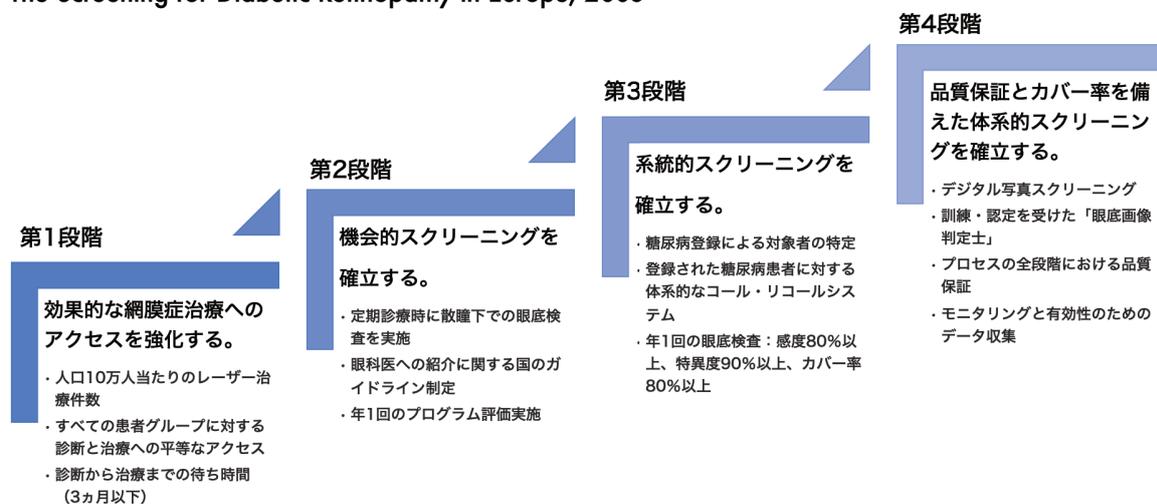
糖尿病網膜症は糖尿病の細小血管合併症として頻度が高く、糖尿病患者の4－5人に一人が罹患していることが報告されている。糖尿病網膜症は高血糖の是正、そして、高血圧、脂質異常症に対する治療といった多面的な内科治療を受けることで発症のリスクを低く抑えることが可能である。糖尿病網膜症に対する治療は大きく進歩し、適切な時期に発見され、適切な治療を受けることでより良い視力を生涯維持することが可能となった。にもかかわらず、いまだに発見が遅れ、適切な治療継続ができない場合には不可逆的な重篤な視力低下や失明に至る症例がある。治療時期を逃すことのないように早期・適時に診断に至る仕組みとしてのスクリーニングの重要性が知られている。そのため、以前から眼底カメラを用いて撮影された眼底写真によるスクリーニングが行われてきた。我が国をはじめ多くの国においては、保険診療の中で内科から眼科への受診勧奨という形で糖尿病網膜症の検診が行われてきた。2018年には特定健康診査において、糖尿病が疑われる対象者に対して糖尿病網膜症の検診を目的に眼底検査を行なっても良いことになったが、十分に普及していない現状がある。

英国をはじめとする欧州各国は1989年に糖尿病の合併症管理に対する考え方の提唱を提唱したSt. Vincent declaration（イタリア・セントビンセント宣言）を踏まえ、単に糖尿病網膜症の検診を個々の患者に対して勧奨することでは十分にスクリーニングが浸透しないことが指摘され、2000年頃から系統だったスクリーニング制度の提唱、そして、実行が進

められてきた。特に、英国では病院外で糖尿病網膜症のスクリーニングを受けることができるNational Health Service (NHS) diabetic eye screening programmeが導入され、12歳以上で、糖尿病の診断を受けた患者は家庭医によってスクリーニングプログラムに登録され、地域の検診センターや検診車において年一回のスクリーニングを受けることができる。このように国をあげて系統的な糖尿病網膜症スクリーニングを提供している国としては、シンガポールや香港がある。系統的なスクリーニングとは、図1に示すような4段階で整備されるべきもので、家庭医（地域住民の健康のために働く総合診療医）から眼科への紹介に頼る経路におけるドロップアウトのリスクを回避できること、また、眼科受診という患者にとっての負担軽減、また、英国では少ない眼科専門医を治療に集中させるといった点のメリットがある。その結果、英国NHSの糖尿病網膜症スクリーニングプログラム（文献3）はその先駆けとなり、国全体で12歳以上の糖尿病患者の82.4%が年一回の糖尿病網膜症スクリーニングを受けることを達成した。（文献4）

ここで、糖尿病網膜症の検出にはデジタル眼底写真撮影という簡便なスクリーニング法が採用されることが多いが、先の通り、2000年代にはいると、様々な医療画像に対する画像処理や病変検出、特徴抽出による自動判定ソフトウェアの開発が盛んになり（文献5）、糖尿病網膜症についてもすでにスクリーニングのために眼底画像が大量に蓄積していたことから、早期から機械学習によるCAD判定、深層学習による診断支援が報告されてきた。2016年には、Googleの深層学習を用いたAIが眼底画像から糖尿病網膜症の診断を高い精度で判別できることが報告され、大きな話題となった。これまでもルールに基づくAI、CADを用いたAIなどがあったが、精度や解析時間において十分ではなかったため、普及するには至らなかったが、深層学習を用いたAIによって、専門医に匹敵する診断能力を持つことが示され、同様のシステムが2018年にはいち早く米国食品医薬品局（FDA）による承認を得るに至っている。その後現在まで、米国では3社の製品が市販されており、また、AIを用いた糖尿病網膜症スクリーニングが保険収載され、実施実績も増加している。我が

## 糖尿病網膜症の系統的スクリーニング構築に向けた4つの段階 The Screening for Diabetic Retinopathy in Europe, 2005



(私訳) [http://www.drscreening2005.org.uk/contact\\_email.html](http://www.drscreening2005.org.uk/contact_email.html) (最終アクセス 2024年6月13日)。

図1

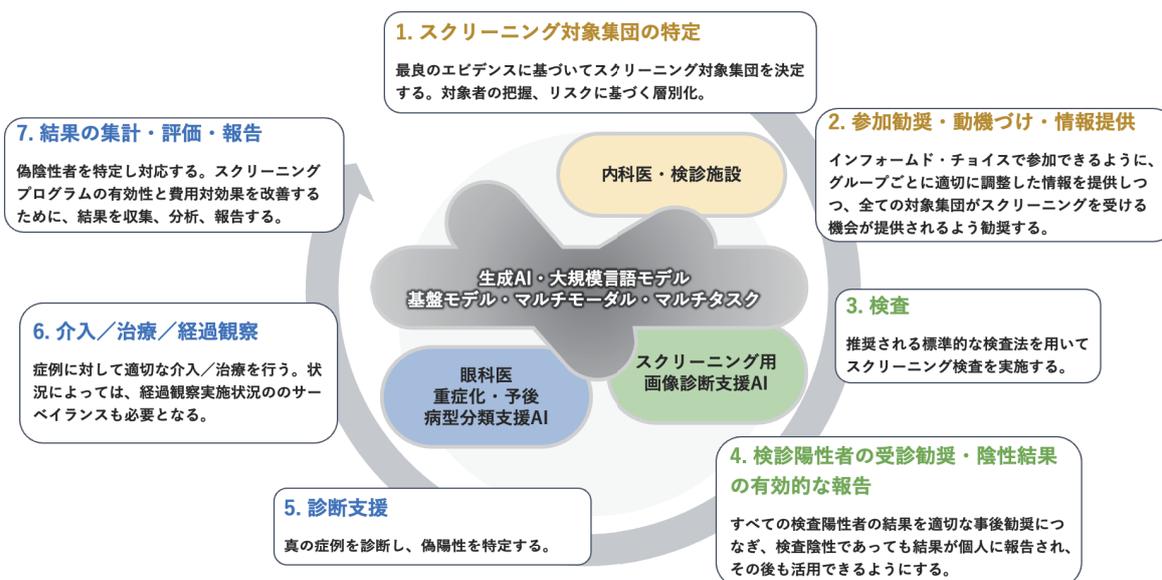
国でも現在2社の医療機器承認済みのAIシステムが販売されている。

深層学習によってもたらされた高い精度の自動診断システムはスクリーニングをさらに発展させる可能性はあるが、一方で、単に従来のスクリーニングに自動診断システムを導入するだけではうまく機能しないという指摘もされている。例えば、タイの11病院において高い精度を達成したGoogle社の糖尿病網膜症自動診断ソフトウェアの実証研究を行った報告（文献6）では、写真撮影から判定結果が得られるまでの時間が2-10週かかっていたところが10分で得られるということから、効率的な健診が期待された。しかし実際には、通信インフラの不安定さ、検診スタッフや患者の人的な要因が影響し、期待した通りの成果が得られないケースがあったという。十分なクオリティの眼底写真が撮影できない場合に判定者が複数の断片的な画像から判定を行うなどの融通が利かない、これまで撮影のみを行い、結果の説明までの時間があったことで、労力が分散されていたものが、すべてをその場で行う必要が出たことでの労務の増

加、また、患者にとっても検診の場で急に眼科受診を勧められるもそれを受け入れられずに嫌がる等问题点が洗い出された。従来のスクリーニングのフローの中で、眼底カメラと診断の部分だけを自動化し、高速化しても必ずしもうまくいかず、スクリーニングのフロー全体をAI診断に合わせて見直し、AI診断ありきで新たなスクリーニングモデルを構築することが求められるだろう。

糖尿病網膜症検出のための眼底写真の判別というAIの応用はすでに医療機器として実装されているが、AI応用の可能性はそれにとどまらず、さまざまな展開を見せている。例えば、画像生成AI技術を応用することで、現在の眼底画像から将来の眼底画像を予測・生成し、その際には、いつ、どこに糖尿病網膜症の所見が出現するかを予測するといったタスクにAIを用いる研究がある。さらにその応用として、これまで糖尿病を有する患者に対して一律に「年一回は症状がなくても眼底スクリーニングを受けましょう」といった勧奨がなされてきたが、予想される網膜症出現時期に合わせて個別化されたスクリーニングイン

## 糖尿病網膜症のスクリーニング・診断・診療の向上にAIはどう活用できるだろうか？



Kawasaki, R. How Can Artificial Intelligence Be Implemented Effectively in Diabetic Retinopathy Screening in Japan? Medicina 2024, 60, 243. を元に演者が日本語版作成

図2

ターバルを提示するといった試みもなされている。さらには、画像予測、画像診断、そして、画像からわかりやすい患者向けの説明文書作成、さらには、検診のリマインダーなど統合的なAI活用の研究事例も報告されている。このように、画像診断・判別支援としての応用から始まった糖尿病網膜症に対するAI活用は、検診と医療の間に、内科と眼科の間に、そして、患者と医師との間に存在する溝に橋をかける役割が期待されている。(図2)(文献7)

### まとめ

医療AIは、ルールに基づくAIという黎明期から始まり、機械学習による萌芽、発展、そして、深層学習、トランスフォーマーといった成熟期に入り、保健医療においても十分な精度を以って応用できるレベルに達している。一方、初期のAIシステムが普及しなかったという事例に学ぶべきは、単に精度の向上など性能が良いだけではなく、それを保健医療分野でどう使うのか、また、人の負担を増やすのではなく、負担を減らし、そして人は人ならではの役割にしっかりと取り組むことができるような「人間中心のAI」のデザインを目指して、実証、応用を進めるべきとの教訓と思われる。

**“The goal is to give to the computer those tasks which it can best do and leave to man that which requires (or seems to require) his judgment. (目標は、コンピュータが最も得意とするタスクをコンピュータに任せ、判断を必要とする (または必要と思われる) タスクを人間に任せることである)”**

*Joseph Weizenbaum (1923-2008)*

### 文献

1. Turing, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, LIX (236):433-460. (<https://doi.org/10.1093/mind/LIX.236.433>)
2. Jones CR, Bergen BK. Large Language Models Pass the Turing Test. arXiv:2503.23674. (<https://doi.org/10.48550/>).
3. Scanlon PH. The English national screening programme for sight-threatening diabetic retinopathy 1. *J Med Screen* 2008;15:1-4.
4. Moreton RBR, Stratton IM, Chave SJ, Lipinski H, Scanlon PH. Factors determining uptake of diabetic retinopathy screening in Oxfordshire. *Diabet Med*. 2017;34 (7):993-999.
5. Nielsen KB, Lautrup ML, Andersen JKH, Savarimuthu TR, Grauslund J. Deep Learning-Based Algorithms in Screening of Diabetic Retinopathy: A Systematic Review of Diagnostic Performance. *Ophthalmol Retina*. 2019;3 (4):294-304.
6. Beede E, Baylor EE, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, Vardoulakis L. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. *Conference on Human Factors in Computing Systems* (<https://dl.acm.org/doi/abs/10.1145/3313831.3376718>).
7. Kawasaki, R. How Can Artificial Intelligence Be Implemented Effectively in Diabetic Retinopathy Screening in Japan? *Medicina* 2024, 60, 243.