

RA 診療における AI への期待

長崎大学 生命医科学域 医療人材連携教育センター

川尻 真也

(2026 年 第 25 回博多リウマチセミナー)

はじめに

関節リウマチ (rheumatoid arthritis : RA) は、慢性炎症を基盤とし、関節破壊や機能障害を来す代表的な自己免疫疾患である。近年、生物学的製剤や分子標的治療薬の進歩により治療成績は大きく向上したが、依然として診断の遅れ、疾患活動性評価のばらつき、治療反応性の個人差といった課題が残されている。

これらの課題に対し、人工知能 (artificial intelligence : AI) をはじめとするデジタル技術は、RA 診療の質と効率を同時に向上させる手段として注目されている。本稿では、AI の基本概念を概説した上で、RA 診療における画像診断、自然言語処理、治療反応性予測への応用、さらに遠隔医療との融合による将来展望について概説する。

1. 人工知能の基礎概念

AI とは、人間の知的活動をコンピュータ上で再現することを目的とした技術の総称である。現在、医療分野で実用化されている AI の多くは、機械学習 (machine learning, ML) およびその一分野であるディープラーニング (deep learning, DL) を基盤としている。

ML は、大量のデータから規則性や判断基準を自動的に学習し、未知のデータに対して分類や予測を行う手法である。RA 診療では、臨床指標、検査値、画像情報などを統合し、疾患分類や予後予測、治療反応性評価に用いられている。

DL は多層のニューラルネットワークを用いることで、特徴量抽出を人手に依存せず自動化できる点に特徴がある。この特性により、関節超音波画像や X 線画像の病変検出、電子カルテ記載内容の解析など、従来は専門家の経験に依存していた作業を高精度に代替・補完することが可能となった。一方で、判断根拠が不透明になりやすい「ブラックボックス性」が課題であり、説明可能 AI の開発が重要視されている。

2. 画像診断支援における AI

RA では関節の炎症や破壊の評価に画像検査（超音波、X 線、MRI）が用いられる。近年では、これらの画像を AI が自動的に解析・評価する技術が急速に進歩しており、画像評価の自動化にディープラーニングが活用され、専門医レベルの精度が報告されている（表 1）。

超音波画像では、RA 滑膜炎の重症度スコア（OMERACT-EULAR 基準の 0~3）を ResNet 系畳み込みニューラルネットワーク（Convolutional neural network, CNN）で自動判定する試みがある¹⁾。静止画だけでなく動画も含めて AI に学習させることで、静止画・動画それぞれにモデルを訓練し評価した結果、AUC 0.74~0.95 程度の判別性能が得られ、モデルの出力と専門医による評価との一致度も高かった（ICC 0.24~0.76）¹⁾。特に動的超音波画像を用いたモデルは静止画モデルよりも精度が高く、熟練放射線科医よりも正確にスコアリングできる場合があると報告された。このような DL モデルにより関節炎の客観的定量評価が可能となり、画像検査の標準化や読影者間差の軽減が期待される。

X 線画像に対しても、関節破壊評価の自動化が進んでいる。RA の手足 X 線における Sharp-van der Heijde（SvH）スコア算定について、国際的なコンペティションである RA2-DREAM チャレンジを通じて高度なモデルが開発された²⁾。その一つである自動 RA スコアリングアルゴリズム（AuRA）アルゴリズムは、まず関節の位置を自動で検出し（YOLOv3）、その後に関節の破壊の程度を評価する（DenseNet）という多段構成をとる²⁾。367 例の X 線画像で訓練後、205 例の外部データで検証したところ、総 SvH スコアの RMSE（RMSE は予測誤差の大きさを示し、値が小さいほどモデル性能が高い）は 23.6 と先行モデル（RMSE 35 前後）を大きく上回る精度を示し、関節破壊進行の予測変化量も専門医評価と高い相関（ $r=0.74$ 、 $p<0.001$ ）を示した。さらに同モデルは数年単位の経時的変化も検出可能であり、X 線による関節破壊進行モニタリングへの応用可能性が示唆されている。また、別の研究では、Vision Transformer（ViT）（画像を「関節という単語の集合」として捉え、全体の文脈から RA の構造的障害を評価する AI）を組み込んだマルチステージモデル（手画像の UNet セグメンテーション+YOLOv7 検出+ViT 予測）によって Sharp スコアの自動推定に取り組み、外部検証で RMSE 9.7、ICC=0.702 を達成したと報告されている³⁾。これは特に軽度~中等度の関節破壊において専門医スコアと良好に一致し、時間と労力を要する SvH 評価の自動化・効率化につながる成果である。

MRI 画像についても、深層学習による病変検出・スコアリングの研究がみられる。例えば 2024 年の研究では、RA 患者の手関節 MRI における骨びらん・骨髄浮腫/骨炎・滑膜炎を CNN で自動分類し、従来の専門医評価と同等の精度で重症度グレードを判定することに成功している⁴⁾。この手法では造影を省略しても高精度な評価が可能であり、撮像プロトコル簡略化による患者負担軽減や読影時間の短縮が期待される。

こうした画像診断支援 AI は、熟練医の少ない施設でも客観的な関節評価を可能にし、経時変化の検出など治療効果判定の精度向上に寄与する可能性がある。ただし、画像データセットの偏りや機器間差による限界も指摘されており、汎用性確保のためより多様なデータでの検証が必要である。

表 1. 画像診断支援に関する主要研究

対象データ・目的	AI 手法	主な性能指標・結果	臨床応用上の示唆 (利点・限界)
手関節の超音波画像による滑膜炎スコア判定 ¹⁾	ResNet 系 CNN	AUC=0.87-0.95 (スコア 0~3 分類) ※熟練放射線科医に匹敵する精度	定量・客観的な関節炎評価が可能 動的画像の活用で専門医超の精度
手足 X 線での Sharp スコア自動算定 ²⁾	YOLOv3+ DenseNet (多段 CNN)	総 SvH スコア推定: RMSE=23.6 ※専門医との差誤差を大幅軽減 経時変化検出: Pearson r=0.74	関節破壊の進行モニタリングに有用 外部検証で実証 ※学習データ依存・他施設検証要
手 X 線での関節破壊スコア算定 ³⁾	UNet+YOLOv7+ViT (マルチステージ)	Sharp スコア推定: RMSE=9.73, MAE=5.35 ICC=0.702 (vs 専門医)	Transformer 活用の初報 軽度病変の評価精度向上 手動採点の省力化に貢献
手 MRI での骨びらん・滑膜炎・骨髄炎検出 ⁴⁾	3D- ResNet CNN (推定)	3 病変の重症度を自動分類 ※診断精度良好 (専門医評価と高一致)	造影を省略しても高精度な評価が可能 読影時間の短縮・定量評価の標準化に期待

3. 自然言語処理・大規模言語モデルによる臨床支援

電子カルテの診療記録や医療文書から有用な情報を抽出・活用する目的で、自然言語処理 (natural language processing, NLP) の応用も進んでいる (表 2)。NLP とは、医師が日常的に記載している文章を AI が理解し、情報として整理・解析する技術である。

近年のレビューでは、NLP 技術は非構造化テキストからの疾患検出や経過管理に有望であり、RA や痛風、脊椎関節炎を高精度に識別できるとされる⁵⁾。実際、複数の研究で Transformer 系モデル (BERT や GPT など) を用いて診療録から RA 患者を同定するアルゴリズムが開発されており、90%以上の高い精度で RA を検出できたケースも報告されている⁵⁾。NLP を活用したフェノタイプ自動抽出により、リウマチ疾患の早期発見や症例集積研究の効率化が期待される。

また、診療録中の疾患活動性指標の抽出にも NLP が応用されている。米国リウマチレジストリ RISE のデータを用いた研究では、医師の外来診療録から DAS28 や HAQ スコアなど

RA の疾患活動性や機能評価の記録値を自動抽出するパイプラインが構築された⁶⁾。約 3450 万件の外来診療録を処理し評価した結果、抽出の感度 95%、F1 スコア（感度と適合率の調和平均であり、実臨床での有用性を反映する指標）0.91 と極めて高い精度を示し、構造化データに記録された値との一致率も良好であった（疾患活動性指標：κ値 0.43~0.68、機能評価指標：κ = 0.86-0.98）。さらに、外部機関データでの検証においても F1 スコア 79%と一定の再現性が確認され、施設ごとに書式や記載方法が異なるカルテからもアウトカム指標を集計できる可能性を示した。このような NLP 技術により、日常診療データから治療成績を効率的に集積し、研究や質改善に活かすことができる。

さらに、RA 合併症領域での NLP 活用も報告されている。例えば RA に合併する間質性肺疾患（RA-ILD）において、診療録中の肺機能検査を自動抽出するアルゴリズムが開発された⁷⁾。退役軍人医療システムのテキストデータから努力性肺活量（FVC）の数値をパターンマッチングで検出し、実際の肺活量計測値と比較したところ、抽出値の 95.8%が実測値の±5%以内に一致し、相関係数 $r=0.94$ という極めて高い精度を達成した。この手法により従来は検査データが揃わず解析できなかった症例を含め、3 倍以上の患者について縦断的な肺機能データを取得でき、RA-ILD の経過研究や臨床モニタリングに大きく貢献しうると考えられる。

NLP は文献情報や患者教育にも応用され始めている。大規模言語モデル（LLM）を用いて医学知識を要約・回答する試みでは、GPT 系モデルが患者からの質問（例えばメトトレキサートの副作用）に対し、専門的でありながら理解しやすい形で回答を生成しうることが示された⁵⁾。このように NLP/LLM は医療者の負担軽減や患者支援にも役立つ可能性があるが、生成される情報の正確性を人が確認することが不可欠であり、医療特有の表現の解釈や誤情報の拡散防止といった課題を踏まえた慎重な運用が求められる。

表 2. 自然言語処理に関する主要研究

用途・対象	手法・モデル	主な性能指標	臨床上の意義・補足
疾患自動検出（カルテ記録から RA 患者識別） ⁵⁾	BERT・GPT など Transformer モデル	RA などリウマチ疾患を高精度検出（例：主要モデルは高い再現率・精度）	診断支援・表現型抽出に有用 大量データから迅速に候補患者を特定
アウトカム指標抽出（診療ノート内の DAS28 等） ⁶⁾	ルールベース NLP パイプライン	RA 疾患活動性/機能評価スコア抽出： 内部検証感度 95%、PPV87%、F1=0.91 外部検証 F1=0.79	経時的アウトカム管理を自動化 非構造テキストから実世界データ活用推進
肺機能データ抽出（RA-ILD の FVC 値） ⁷⁾	正規表現マッチング + 前処理	相関 $r=0.94$ （NLP 抽出値 vs 機器測定値） 95.8%が±5%以内の誤差	肺合併症のモニタリング効率化 患者あたり縦断データ点 3 倍以上を取得

用途・対象	手法・モデル	主な性能指標	臨床上の意義・補足
LLMによる患者支援（治療説明・QA） ⁵⁾	GPT系大型言語モデル（ChatGPT等）	患者からの質問に対し高精度で適切な回答を生成（例：メトトレキサートの情報提供）	患者教育や遠隔医療サポートに可能性 ※医療知識の最新性・正確性担保が課題

4. 治療反応性予測におけるAI

患者ごとに異なる治療反応を事前に予測し、最適な治療選択につなげることはRA診療における重要課題である。近年、この治療反応性予測にMLを活用する研究が増えており、バイオマーカーや電子カルテ情報、患者報告アウトカムなど多次元データから治療効果を予測するモデルが数多く報告されている（表3）。

特に生物学的製剤（bDMARD）投与時の治療効果予測に関する研究が活発である。最新のスコopingレビューによれば、RAの治療反応予測にMLを用いた研究24件のモデル性能（ROC曲線下面積：AUC）は0.54～0.92と幅広く、平均は約0.71であった⁸⁾。すなわち、一定の予測精度を示すモデルが存在する一方で、性能にはばらつきがあることが示されている。モデル手法としては、決定木ベースの勾配ブースティング（XGBoostなど）やランダムフォレスト、サポートベクターマシン、DLが多用され、なかでも複数の情報を柔軟に組み合わせられるブーステッドツリー系やニューラルネットが良好な成績を示す傾向にあった。入力変数としては疾患活動性スコア（DAS28など）や炎症マーカー、機能評価、患者報告症状が共通して重要視されており、モデルの解釈性確保のため説明可能AIであるSHAP（SHapley Additive exPlanations）値などでどの因子が予測に強く影響したかを可視化する研究も行われている⁹⁾。一方で外部検証を行っている研究は全体の17.5%に留まり、コホート間での再現性検証や手法の標準化が今後の課題と指摘されている⁸⁾。

具体的な例として、初回bDMARD治療で6か月後に寛解（DAS28<2.6）に到達するかを予測するモデルでは、治療開始時点のルーチン臨床データからXGBoost（臨床・検査・画像スコアのような異種データを統合し、RAの再燃や治療反応を高精度かつ説明可能に予測できる実用的なML手法）によりAUC=0.91という高い精度が報告されている¹⁰⁾。同一コホートで12か月間の寛解持続を予測する際にはAdaBoostモデルが最適とされ、AUC=0.84を達成した。これらのモデルではDAS28、患者VAS、年齢、腫脹関節数などが予測に寄与する重要特徴と同定されており⁹⁾、予測根拠を示すことで臨床応用への信頼性も高めている。予測確率のキャリブレーションも施すことで、モデル出力をリスク層別化に活用し、寛解未達時のフォローアップ間隔調整や治療変更判断に役立つ試みもなされている¹⁰⁾。

また、寛解中患者の再燃リスク予測にMLを応用した研究では、超音波で検出される滑膜血流（SMI法）や炎症マーカーを含む73項目の特徴量から、2年以内の再燃をAUC 0.75で予測できたと報告されている¹¹⁾。この研究ではXGBoostモデルが従来の臨床指標のみの予測より精度が高く、特に手関節および中足趾関節のSMIスコアが重要因子に挙げられた。MLに

より経験的に知られていた因子の重要度を定量化し、新たな組合せ予測因子の発見につながる点でも意義がある。

総じて、RA の治療反応予測モデルは一定の有望性を示しているものの、実臨床で実用化するにはさらなる検証が必要である。多数の研究が単施設・既存データセットで行われており、異なる集団への適用可能性や汎用性については不明な点が多い。今後、モデルの外部検証や国際共同研究による評価、ならびに予測結果の解釈性向上や意思決定支援ツールとしての検証が進めば、個別化医療の一環としてこれら AI モデルが治療戦略の最適化に寄与することが期待される。

表 3. 治療反応性予測に関する主要研究

研究内容・対象	モデル・手法	性能指標（予測精度）	臨床応用上のポイント
RA 治療反応予測（24 研究分析） ⁸⁾	RF, XGB, SVM, NN など 各種 ML 手法	AUC 範囲 0.54～0.92（平均 0.71） ※外部検証は 17.5%のみ	モデル間で性能ばらつき 高性能例もあり（AUC~0.8-0.9） ※標準化・再現性が課題
6 か月、12 か月寛解予測（bDMARD） ¹⁰⁾	XGBoost、AdaBoost （ブーステッド決定木）	6 か月寛解達成の予測：AUC=0.91 12 か月寛解持続の予測：AUC=0.84	高精度モデルの実現 確率出力を校正・リスク分類 ※中止判断やフォロー計画に活用提案
再燃予測（寛解患者） ¹¹⁾	XGBoost （臨床+US+血液データ）	AUC=0.747（2 年以内再燃） RF: 0.719, LR: 0.701	複数バイオマーカー統合で精度向上 重要因子: 手関節/MTP の SMI スコア等

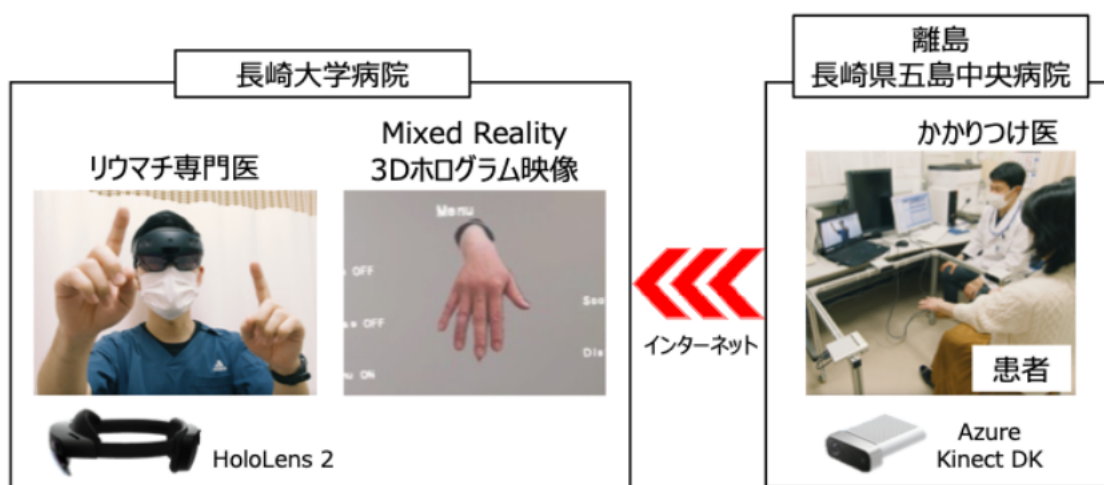
以上のように、RA 診療の各領域で AI 技術の応用が進みつつある。画像解析 AI は客観的な疾患評価やモニタリングを支援し、NLP は日常診療データから有益な知見を引き出し、ML 予測モデルは治療戦略の個別化に寄与し得る。今後さらなるエビデンスの蓄積と技術改良により、関節リウマチ診療における AI は診断・治療の精度向上と効率化に大きく貢献することが期待される。

5. AI と遠隔医療が拓く RA 診療の未来

近年、著者らは Mixed Reality 技術および AI を活用した遠隔医療プラットフォームである NURAS (Nagasaki University Rheumatoid Arthritis Remote Medical System、図 1) を開発し、RA 診療への応用を進めている。NURAS は、関節診察情報や関節超音波画像を三次元的に可視化・共有することで、専門医が遠隔地から診療支援を行うことを可能とするシステムである¹²⁾。本システムにより、従来の二次元画像共有では把握が困難であった関節の空間的情報を遠隔環境でも評価でき、診療の質向上が期待される。また、AI による関節超音波画像の自動スコアリングや診断支援と統合することで、疾患活動性評価の標準化や診療効率の向上に寄与しうる。

さらに NURAS は、専門医の診察視点や評価プロセスを共有できる点で教育的意義も大きく、非専門医や若手医師への指導ツールとしての活用も期待される。一方で、通信環境や情報セキュリティ、診療責任の整理といった課題も残されており、AI や遠隔医療は診療を補完する技術であるという前提を踏まえた運用が重要である。

図 1. NURAS の概要



6. おわりに

AI は RA 診療において、診断、評価、治療選択、さらには遠隔医療を含む診療体制の変革をもたらす可能性を有している。AI は医師に取って代わるものではなく、診療を補完し、意思決定を支援するツールである。今後、臨床現場での検証と社会実装を進めることで、AI は RA 診療の質向上と医療格差是正に大きく貢献すると期待される。

文献

- 1) He X, Wang M, Zhao C, et al. Deep learning-based automatic scoring models for the disease activity of rheumatoid arthritis based on multimodal ultrasound images. *Rheumatology (Oxford)*. 2024;63:866-873.
- 2) Venäläinen MS, Biehl A, Holstila M, Kuusalo L, Elo LL. Deep learning enables automatic detection of joint damage progression in rheumatoid arthritis-model development and external validation. *Rheumatology (Oxford)*. 2025;64:1068-1076.
- 3) Moradmamand H, Ren L. Multistage deep learning methods for automating radiographic sharp score prediction in rheumatoid arthritis. *Sci Rep*. 2025;15:3391.
- 4) Schlereth M, Mutlu MY, Utz J, et al. Deep learning-based classification of erosion, synovitis and osteitis in hand MRI of patients with inflammatory arthritis. *RMD Open*. 2024;10:e004273.
- 5) Omar M, Naffaa ME, Glicksberg BS, Reuveni H, Nadkarni GN, Klang E. Advancing rheumatology with natural language processing: insights and prospects from a systematic review. *Rheumatol Adv Pract*. 2024;8(4):rkae120.
- 6) Humbert-Droz M, Izadi Z, Schmajuk G, Gianfrancesco M, Baker MC, Yazdany J, Tamang S. Development of a Natural Language Processing System for Extracting Rheumatoid Arthritis Outcomes From Clinical Notes Using the National Rheumatology Informatics System for Effectiveness Registry. *Arthritis Care Res (Hoboken)*. 2023;75:608-615.
- 7) England BR, Roul P, Yang Y, et al. Extracting forced vital capacity from the electronic health record through natural language processing in rheumatoid arthritis-associated interstitial lung disease. *Pharmacoepidemiol Drug Saf*. 2024;33:e5744.
- 8) Eriakha EB, Han Y, Li M, Li J, Huang Y. Machine learning for predicting treatment response to biologic and targeted synthetic disease-modifying antirheumatic drugs in rheumatoid arthritis: a scoping review. *BMC Rheumatol*. 2025;9:132.
- 9) Salehi F, Salin E, Smarr B, et al. A robust machine learning approach to predicting remission and stratifying risk in rheumatoid arthritis patients treated with bDMARDs. *Sci Rep*. 2025;15:23960.
- 10) Salehi F, Lopera Gonzalez LI, Bayat S, et al. Machine Learning Prediction of Treatment Response to Biological Disease-Modifying Antirheumatic Drugs in Rheumatoid Arthritis. *J Clin Med*. 2024;13:3890.
- 11) Matsuo H, Kamada M, Imamura A, et al. Machine learning-based prediction of relapse in rheumatoid arthritis patients using data on ultrasound examination and blood test. *Sci Rep*. 2022;12:7224.
- 12) 令和 6 年度厚生労働省科学研究費補助金（免疫・アレルギー疾患政策研究事業）研究。患者視点に立ったリウマチ疾患のアンメットメディカルニーズの「見える」化と社会実装に資する研究班。関節リウマチ、若年性特発性関節炎（少関節炎型・多関節炎型）、全身性エリテマトーデスの遠隔医療の確立に向けた提言。 https://www.ryumachi-jp.com/jcr_wp/media/2025/05/news250601_recommendation.pdf