

令和 6 年 6 月 6 日現在

機関番号：32203

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K10355

研究課題名（和文）自動問診AIの利用を前提とした診断精度を最大化する診断思考理論の開発

研究課題名（英文）Developing a diagnostic thinking strategy to maximize the diagnostic accuracy when using AI-based automated medical history taking systems

研究代表者

原田 侑典（Harada, Yukinori）

獨協医科大学・医学部・講師

研究者番号：40810502

交付決定額（研究期間全体）：（直接経費） 2,300,000円

研究成果の概要（和文）：本研究ではAI自動問診による鑑別診断リストの提示の有無は医師の診断精度に影響を与えないが、AIの鑑別診断リストの診断精度は医師の診断精度に影響を及ぼすことを示した。次に、一般内科外来でAI自動問診を利用した場合、診断エラーの頻度が若干低下する可能性を示した。さらに、AI自動問診の鑑別診断リストの診断精度は、他のAIの鑑別診断リストと共通する鑑別診断の数が多いほど高いことを示し、複数のAIを併用することに有用性を示唆した。また、AI自動問診の診断精度は経時的に向上しておらず、稀な疾患や病像では精度が低い問題も指摘した。さらに、AI自動問診の診断精度を医師は正確に判断できないことも示した。

研究成果の学術的意義や社会的意義

本研究の結果から、現在日本で利用されているAI自動問診の診断精度は全面的に信用してよいほどの水準にはないこと、特に稀な疾患や非典型的な病像など診断に誤りや遅れが生じる危険性が高い患者においてはさらに低くなることから、そのような患者であると感じた場合にはAIを頼らない方がよいことが示唆されるほか、AIの診断が正しいか否かの判断は医師の直観は当てにはならないため、他のAIを併用して共通する鑑別が多いかどうかで判断する方が安全であることも示唆される。このように、本研究はAI自動問診を安全に使用する際の具体的な方法を推奨することができた。

研究成果の概要（英文）：In our study, we demonstrated that while the presence of an AI-generated differential diagnosis list does not inherently influence physicians' diagnostic accuracy, the accuracy of the AI's diagnoses does. Additionally, we found that the use of AI for medical history taking in general internal medicine outpatient clinics could marginally decrease diagnostic errors. The diagnostic precision of AI improved with the increase in shared diagnoses across different AI systems, highlighting the benefit of utilizing multiple AI tools. However, we observed that the diagnostic capability of AI has not advanced over time and remains limited for uncommon diseases and atypical presentations. Moreover, we showed that physicians cannot correctly assess the reliability of diagnoses provided by AI-driven medical history taking systems.

研究分野：診断エラー

キーワード：AI自動問診 診断エラー

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

問診は診断において重要な役割を担うが、医師は診断に必須な質問の 6 割しか聞くことができていないとされる。特に、経験や知識が不足した医師はそもそも鑑別診断を想起することができないため、診断につながる重要な情報を問診で聞き取ることができず、診断精度が落ちる。鑑別診断の想起を促す仕組みを取り入れると問診および診断の精度が上がることは示されているが、医療現場では患者の個性が高く、一般化が困難であるため普及していない。患者の個性、および人間の判断・認知・意思決定に紛れ込む恣意的な傾向である認知バイアスの影響を受けずに鑑別疾患想起と適切な問診を補助する策の 1 つに人工知能 (AI) 技術の応用がある。近年、AI が自動で患者の問診を行い、予測される鑑別診断を挙げる自動問診 AI が医療現場に導入されるようになった。先行研究では自動問診 AI の診断精度は症例の複雑度などによって変化することが知られており、AI を利用する際には、AI を信用して良い状況と良くない状況を判断する必要があることが示唆されていたが、医師にそのような判断ができるのかは明確にはなっていない。このことは、医師が AI を有効利用できていないばかりか、AI の利用によって診断の誤り、遅れ、見逃しといった診断エラーが生じることさえある可能性を示唆していた。よって、自動問診 AI 技術を活用して医師の診断精度を上げるためには、自動問診 AI を最大活用するために最適な医師の診断思考過程を解明する必要がある。

2. 研究の目的

本研究の目的は、自動問診 AI の状況別の詳細な診断精度、自動問診 AI を利用した医師の診断精度に影響を与える要因、自動問診 AI を利用した際に起こる診断エラーの原因と対策を検証し、自動問診 AI を利用した医師の診断精度を最大とする診断思考方法を理論化することであった。

3. 研究の方法

(1) 自動問診 AI の鑑別診断リストの提示が医師の診断に及ぼす影響の検証

内科外来を受診し、24 時間以内に入院した患者 16 人分の自動問診 AI が作成した病歴と鑑別診断リストを利用し、獨協医科大学病院総合診療科または臨床研修センターに所属する医師 22 人を対象にランダム化比較試験を行った。参加医師を介入群または対照群にランダムに割り当て、介入群は自動問診 AI が作成した病歴と鑑別診断リストを読み、対照群は自動問診 AI が作成した病歴のみを読んだ。両群ともに参加医師は 1-3 個の鑑別診断を可能性の高い順に記載し、医師の鑑別診断の中に最終診断が含まれている割合を 2 群で比較した。また、医師の鑑別診断の中に最終診断が含まれていることと関連する要因について多因子ロジスティック回帰モデルを利用して解析した。

(2) 内科外来で自動問診 AI を利用した患者における診断エラー発生率の検証

内科外来を受診して自動問診 AI を利用し、その後 14 日以内に予期せぬ入院に至った患者 146 人を対象にした観察研究を行った。診断エラーが起きたかどうかを判定するツールである The Revised Safer Dx Instrument を用いて対象患者の診療録を評価し、診断エラー発生率を検証した。また、自動問診 AI の鑑別診断リストの中に最終診断が含まれていた患者と含まれていなかった患者における診断エラー発生率を比較した。さらに、診断エラーありと判定された患者について背景要因の検討を行った。

(3) 他の鑑別診断リスト生成 AI の利用による自動問診 AI の鑑別診断リストの精度の予測

自動問診 AI が作成した 103 人分の患者の病歴を利用し、研究対象の自動問診 AI とは異なる 2 つの AI を用いてそれぞれ鑑別診断リストを生成し、同じ症例に対して作成された 3 つの鑑別診断リストの中に共通する診断の個数と、自動問診 AI の鑑別診断リストに最終診断が含まれている割合の関係を調べた。

(4) 自動問診 AI の鑑別診断リストの診断精度が経時的に向上しているか否かの検証

2019 年度から 2022 年度にかけての 3 年間に、内科外来を受診して自動問診 AI を利用し、その後 30 日以内に入院した患者 381 人を対象として、1 年毎の自動問診 AI の鑑別診断リストの診断精度を比較した。また、疾患を頻度の観点から 2 つ (頻度の高い疾患と頻度の低い疾患)、病歴を典型度の観点から 2 つ (典型と非典型) に分けて、それぞれのグループにおける自動問診 AI の鑑別診断リストの診断精度も検証した。

(5) 自動問診 AI の鑑別診断リストの診断精度に対する信頼校正によって医師の診断精度が変わるか

内科外来を受診し、30 日以内に入院した患者 20 人分の自動問診 AI が作成した病歴と鑑別診断リストを利用し、獨協医科大学病院総合診療科に所属したことがある医師 20 人を対象に 2 群間比較試験を行った。参加医師を介入群または対照群に割り当てた。両群ともに自動問診 AI が作成した病歴と鑑別診断リストを読み、1-3 個の鑑別診断を可能性の高い順に記載した。介入群では「自動問診 AI の鑑別疾患リストに正しい診断が含まれているかを検討してください」という

問いかけを追加した。この問いかけを行うことで自動問診 AI の診断精度に対する信頼較正を誘導することを試みた。医師の鑑別診断の中に最終診断が含まれている割合を 2 群で比較した。

4. 研究成果

(1) 自動問診 AI の鑑別診断リストの提示が医師の診断に及ぼす影響の検証

自動問診 AI の鑑別診断リストの提示の有無によって医師の診断精度に差はなかった（正答率：介入群 57.4%、対照群 56.3%、 $P=0.91$ ）。AI の鑑別診断リストに最終診断が含まれている場合、医師の診断正答率は高かった（調整オッズ比 7.7、95%信頼区間 4.7-12.6）。医師の診断に AI の鑑別診断リストの診断が含まれている割合は介入群 70.2%、対照群 55.1%と介入群で有意に高かった（ $P<0.001$ ）。介入群では、AI の鑑別診断リストに最終診断が含まれているが医師の鑑別診断リストに最終診断が含まれていない割合が 15.9%、AI の鑑別診断リストに最終診断が含まれていないが医師の鑑別診断リストの全てに AI の鑑別診断リストの診断が記載されている割合は 14.8%であった。

(2) 内科外来で自動問診 AI を利用した患者における診断エラー発生率の検証

対象患者 146 人のうち 138 人で最終診断が判明した。138 人のうち最終診断が AI の鑑別診断リストに含まれていたのは 69 人（50%）であった。診断エラーは 146 人中 16 人に認められた（11.0%、95%信頼区間 6.4%-17.2%）。診断エラーの発生率に関して、AI の鑑別診断リストに最終診断が含まれている群では含まれていない群と比較して低い傾向があったが、統計学的有意な差は認めなかった（7.2% vs 15.9%、 $P=0.18$ ）。

(3) 他の鑑別診断リスト生成 AI の利用による自動問診 AI の鑑別診断リストの精度の予測

自動問診 AI の鑑別診断リストのみの診断正答率は 46%であった。自動問診 AI が作成した鑑別診断リストと他の AI によって生成された鑑別診断リストの中に一致する診断の数が増えるほど、自動問診 AI の鑑別診断リストの中に最終診断が含まれている可能性は高くなった（他の 1 つの AI が生成した鑑別診断リストと同一診断がない場合は、自動問診 AI の鑑別診断リストに最終診断が含まれている割合は 20%に対して、5 つの診断が一致している場合は 78%まで上昇した、診断が一致する数が 1 増えるごとに調整オッズ比 1.5、95%信頼区間 1.3-1.7）（図 1）。

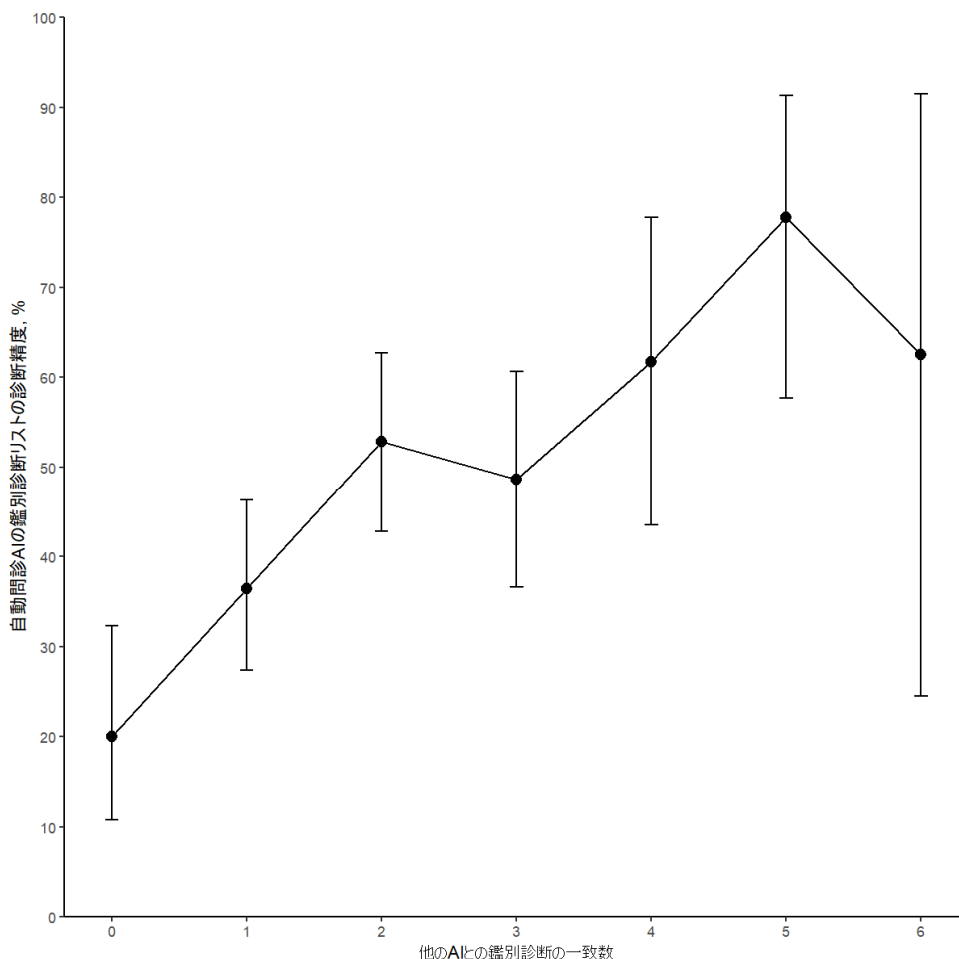


図 1. 自動問診 AI の鑑別診断リストと他の鑑別診断リスト生成 AI の鑑別診断リストの鑑別診断の一致数と自動問診 AI の鑑別診断リストの診断精度の関係

(4) 自動問診 AI の鑑別診断リストの診断精度が経時的に向上しているか否かの検証

自動問診 AI の鑑別診断リストに最終診断が含まれている割合は 45%であり、調査した 3 年間で経時的な変化は認めなかった (1 年目 44%、2 年目 44%、3 年目 48%、 $P=0.85$) (表 1)。自動問診 AI の鑑別診断リストに最終診断が含まれている割合は、頻度の低い疾患と比較して頻度の高い疾患では高く (オッズ比 4.1、95%信頼区間 2.5-7.0)、非典型的な病歴と比較して典型的な病歴で高くなった (オッズ比 6.9、95%信頼区間 3.6-14.2)。

表 1. 自動問診 AI の鑑別診断リストの診断精度の 3 年間の変化

	全体 N=381	1 年目 N=219	2 年目 N=72	3 年目 N=90	P 値
全体の診断精度	172/381 (45%)	97/219 (44%)	32/72 (44%)	43/90 (48%)	0.85
頻度の高い疾患	142/257 (55%)	79/146 (54%)	27/45 (60%)	36/66 (55%)	0.78
頻度の低い疾患	30/124 (24%)	18/73 (25%)	5/27 (19%)	7/24 (29%)	0.67
典型的な病歴	160/298 (54%)	88/164 (54%)	29/60 (48%)	43/74 (58%)	0.53
非典型的な病歴	12/83 (15%)	9/55 (16%)	3/12 (25%)	0/16 (0%)	0.10

(5) 自動問診 AI の鑑別診断リストの診断精度に対する信頼校正によって医師の診断精度が変わるか

「自動問診 AI の鑑別疾患リストに正しい診断が含まれているかを検討してください」という問いかけの有無で医師の診断正答率に差はなかった (介入群 41.5%、対照群 46.0%、 $P=0.42$)。介入群で自動問診 AI の鑑別疾患リストに最終診断が含まれているか否かの判定が正しかったのは 61.5%のみであった。

上記 5 つの研究結果から、自動問診 AI の診断精度は必ずしも高いとは言えず、経時的にも改善しているとは言えないこと、自動問診 AI は頻度の低い疾患や非典型的な病歴を呈する患者における診断精度が下がること、自動問診 AI の鑑別診断リストを提示することで医師の鑑別診断の内容に影響は出るものの診断精度自体には影響が出ないこと、自動問診 AI の鑑別診断リストの精度についての医師の予測精度は高くないこと、しかしながら、自動問診 AI 利用後の診断エラー発生率は低くなる可能性があることが示された。これらの知見から、自動問診 AI を利用した医師の診断精度を最大とするためには、自動問診 AI の鑑別診断リストに最終診断が含まれている割合は 5 割程度であることを意識し、他の AI で作成した鑑別診断リストとの鑑別診断一致数が多ければ信用度を上昇させ、鑑別診断一致数が少なければ信用度が低下させた上で、信用度が高いと判断した場合には自動問診 AI の鑑別診断リストを参考にし、信用度が低いと判断した場合には病歴だけを参考にして自身で改めて鑑別診断を考える戦略が有効であると考えられた。これらの成果は最近の生成 AI についても応用可能だと考えられる。

5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 5件/うち国際共著 0件/うちオープンアクセス 5件）

1. 著者名 Harada Yukinori, Sakamoto Tetsu, Sugimoto Shu, Shimizu Taro	4. 巻 -
2. 論文標題 Longitudinal changes in diagnostic accuracy of a differential diagnosis list developed by an artificial intelligence-based symptom checker: a retrospective observational study	5. 発行年 2024年
3. 雑誌名 JMIR Formative Research	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.2196/53985	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Harada Yukinori, Tomiyama Shusaku, Sakamoto Tetsu, Sugimoto Shu, Kawamura Ren, Yokose Masashi, Hayashi Arisa, Shimizu Taro	4. 巻 7
2. 論文標題 Effects of Combinational Use of Additional Differential Diagnostic Generators on the Diagnostic Accuracy of the Differential Diagnosis List Developed by an Artificial Intelligence?Driven Automated History?Taking System: Pilot Cross-Sectional Study	5. 発行年 2023年
3. 雑誌名 JMIR Formative Research	6. 最初と最後の頁 e49034 ~ e49034
掲載論文のDOI（デジタルオブジェクト識別子） 10.2196/49034	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Harada Yukinori, Katsukura Shinichi, Kawamura Ren, Shimizu Taro	4. 巻 18
2. 論文標題 Efficacy of Artificial-Intelligence-Driven Differential-Diagnosis List on the Diagnostic Accuracy of Physicians: An Open-Label Randomized Controlled Study	5. 発行年 2021年
3. 雑誌名 International Journal of Environmental Research and Public Health	6. 最初と最後の頁 2086 ~ 2086
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/ijerph18042086	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Harada Yukinori, Katsukura Shinichi, Kawamura Ren, Shimizu Taro	4. 巻 18
2. 論文標題 Effects of a Differential Diagnosis List of Artificial Intelligence on Differential Diagnoses by Physicians: An Exploratory Analysis of Data from a Randomized Controlled Study	5. 発行年 2021年
3. 雑誌名 International Journal of Environmental Research and Public Health	6. 最初と最後の頁 5562 ~ 5562
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/ijerph18115562	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Kawamura Ren, Harada Yukinori, Sugimoto Shu, Nagase Yuichiro, Katsukura Shinichi, Shimizu Taro	4. 巻 10
2. 論文標題 Incidence of Diagnostic Errors Among Unexpectedly Hospitalized Patients Using an Automated Medical History-Taking System With a Differential Diagnosis Generator: Retrospective Observational Study	5. 発行年 2022年
3. 雑誌名 JMIR Medical Informatics	6. 最初と最後の頁 e35225 ~ e35225
掲載論文のDOI (デジタルオブジェクト識別子) 10.2196/35225	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計4件 (うち招待講演 0件 / うち国際学会 2件)

1. 発表者名 Tetsu Sakamoto, Yukinori Harada, Taro Shimizu
2. 発表標題 Does facilitating trust calibration for artificial-intelligence-driven differential diagnosis list improve physicians' diagnostic accuracy?: a quasi-experimental study
3. 学会等名 ACP (米国内科学会) 日本支部年次総会・講演会2024
4. 発表年 2024年

1. 発表者名 Harada Y, Tomiyama S, Sakamoto T, Yokose M, Kawamura R, Hayashi A, Shimizu T
2. 発表標題 Accuracy of Ddx generator based on the number of shared Ddx with other Ddx generators
3. 学会等名 SIDM 2023 Cleveland (国際学会)
4. 発表年 2023年

1. 発表者名 Arisa Hayashi, Yukinori Harada, Shusaku Tomiyama, Masashi Yokose, Tetsu Sakamoto, Ren Kawamura, Taro Shimizu
2. 発表標題 Improvement of diagnostic accuracy of artificial intelligence-based differential diagnosis lists by using multiple artificial intelligence systems
3. 学会等名 SIDM 's 15th Annual Society to Improve Diagnosis in Medicine conference, SIDM 2022 (国際学会)
4. 発表年 2022年

1. 発表者名 Katsukura S, Harada Y, Kawamura R, Shimizu T
2. 発表標題 The AI-driven differential diagnosis list was not associated with the diagnostic accuracy of physicians when using the AI-driven automated medical history taking systems: a randomized trial
3. 学会等名 ACP(米国内科学会)日本支部年次総会・講演会2021
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	志水 太郎 (Shimizu Taro) (50810529)	獨協医科大学・医学部・教授 (32203)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------