

ICT を利用した医師国家試験の評価方法の開発と検証のための研究

研究代表者	河北 博文	(公益財団法人日本医療機能評価機構 理事長)
研究分担者	浅田 義和	(自治医科大学医学教育センター 准教授)
	岡崎 仁昭	(自治医科大学医学教育センター 顧問)
	久保 沙織	(東北大学教育学研究科 准教授)
	小西 靖彦	(順天堂大学医学教育研究室 特任教授)
	小松 弘幸	(宮崎大学医学部 医療人育成推進センター 教授)
	高村 昭輝	(富山大学学術研究部医学系 医学教育学 教授)
	内藤 俊夫	(順天堂大学大学院医学研究科 教授)
	奈良 信雄	(順天堂大学医学部 客員教授)
	錦織 宏	(名古屋大学医学系研究科 教授)
	伴 信太郎	(愛知医科大学医学教育センター 特命教育教授)
	藤川 裕恭	(順天堂大学医学部総合診療科学講座 特任助教)
	松山 泰	(自治医科大学医学教育センター 教授)
	森 博威	(順天堂大学大学院医学研究科 准教授)

研究要旨

我が国は、2040 年には少子高齢化を超えて高齢者がピークを迎える時代を控え、同時に AI による診断・治療技術の進歩等が予想されている。その中で、人口減少に対する医師数や臓器別専門医、総合診療等の役割、地域医療教育など、医学教育に関して様々な課題がある。令和4年に改訂された医学教育モデル・コア・カリキュラムは、2040 年に活躍する医師を想定して「総合的に患者・生活者をみる姿勢」と「情報・科学技術を活かす能力」を医学生が身につけるべき新たな資質・能力として定義した。また、患者の医療に関するリテラシーが大きく変わることが予見され、受容・傾聴・共感など、診療の中での医師の対話力や姿勢は医師の資質として一層強く求められる。さらに、AI の普及に伴い、AI を活用した医療においてその責任を持つ医師が求められる。このため、将来のあるべき医師像を見据えて、医学生の学修を促進する評価の軸を定めることは極めて重要である。

本研究は、「(1) CBT 医師国家試験の導入の実現に向けた課題の抽出」「(2) 医学生の技能や態度の評価」「(3) 医師国家試験出題基準の検証」「(4) 外国の医学部を卒業した者に対する予備試験および日本語診療能力調査の課題」に関してそれぞれの研究を一体的に行い、我が国のより良い医学・医療に貢献できる医学生の評価軸を示すとともに、医学教育、特に臨床実習の充実につながる医師国家試験の CBT 化の実現に結びつけることを目標とした。そこで、国において医師国家試験と医学教育に関する検討が行われる際の資料として活用できる知見を得るとともに、今後の医師国家試験と医学教育の改善に向けて提言することを目的とする。

(1) CBT 医師国家試験の導入の実現に向けた研究では、CBT 医師国家試験の導入の実現に向けた課題等を抽出するために、前年度に引き続いて、インターネットを介した CBT 試験システム(TAO)を利用して、動画・音声等を取り入れた実臨床に近い問題を取り入れて、全国60大学(医学生 1,647名)の参加による CBT トライアル試験を、大学会場形式と一般会場形式で行い、どちらも大きな問題なく概ね円滑に実施し、将来の試験運営における知見を整理することができた。CBT 問題の作問・ライブラリシステムについては、問題の整理・検索の機能を強化するための検討を行った結果、医師国家試験出題基準等に準じたプール問題の整理・検索が可能となった。CBT 問題の体系的な管理基盤を

MooTAO 上で構築するとともに、カスタム GPT により臨床問題の派生問題等を短期間に多数作成できることを実証した。IRT に基づくテスト運用では、等化の手続きを経た項目の集まりである項目プールをあらかじめ用意する必要があることから、等化係数法による項目母数の等化の一例を示した。医師国家試験の CBT 導入において、IRT を用いた CBT 形式により、均質な問題セットを作成して試験問題の質担保を図るために、試験問題の非公開が必須である。

(2) 医学生の実能や態度の評価に関する研究では、ワークショップを通じて、日本における総括的 Workplace-Based Assessment (WBA) 導入に関する主要論点を整理し、その知見を踏まえて、我が国の臨床実習における総括的 WBA の実施指針案として Genba Based Assessment (GenBA) を作成した。さらに、障害者支援および医療安全の専門家との意見交換を通じて、本質的能力の定義、合理的配慮、患者安全を踏まえた評価基準設定など、制度設計上の重要課題を明らかにした。

(3) 医師国家試験出題基準に関する研究では、令和 6 年版医師国家試験出題基準の医学各論において、延べ 476 項目の備考欄への記載が認められ、その内訳には追加病名が最も多かった。研究者による探索的検討では、延べ 476 項目のうち 344 項目が削除可能な項目、132 項目が検討を要する項目と判断された。今後、専門家等による妥当性検証を行い、備考欄の役割を明確化したうえで、出題基準改定時の整理方針を検討する必要がある。医学総論および必修の基本的事項の出題領域別の検証については、レベル I 問題の出題割合が 50% 以上であった項目は、必修の基本的事項は、「1. 医師のプロフェッショナリズム」「3. 診療情報と諸証明書」等の 5 項目であり、医学総論は、「Ⅰ. 保健医療論」「Ⅱ. 予防と健康管理・増進」等の 3 項目であった。これらについて、ブループリントにおける出題割合を減らし、その分を臨床実習開始前に実施される共用試験 CBT で一定程度出題するなどの調整が望ましいことを示した。

(4) 外国の医学部を卒業した者に対する予備試験および日本語診療能力調査に関する研究では、外国の医学部卒業生の受け入れ、勤務状況に関しては大きな問題は見られなかった。一方、少数ではあるが外国の医学部出身医師の勤務状況に問題が認められたこと、および今後も外国の医学部出身医師の増加が見込まれることから、国民から信頼が得られるように、さらに詳細な調査研究を行い、より適正な外国の医学部出身医師の受け入れ制度を構築する必要性を示した。

今後、国において医師国家試験と医学教育について審議が行われる場合には、本研究で得られた成果を踏まえて議論されることを望む。さらに、国の施策として医師国家試験の CBT 化が決定された場合は、本研究で培った知見やノウハウが活用されることを期待する。

A. 研究目的

我が国の医療の質の向上を図るためには、卒前教育、卒後教育、生涯教育のシームレスな医学教育の取り組みの充実が重要である。2040 年には AI による診断技術の進歩や患者の医療に関するリテラシーが大きく変わることが予見され、受容・傾聴・共感など、診察の中での医師の対話力や姿勢は医師の力量として一層強く求められる。言い換えれば、ヒューマンタッチのできる医師が大切である。これらに対応できる将来のあるべき医師像を見据えて、国民から信頼される臨床能力に優れた

医師を養成することが求められる。

現在、我が国における医師国家試験は冊子形式での多選択肢問題 (Multiple Choice Question : MCQ) で行われているが、MCQ は「知識」の評価には適しているものの、知識と同様に重要である、「技能」や「態度」の評価には十分ではない。

現在、各大学で行われている臨床実習における学生の評価、Post-CC OSCE、および PBT (Paper-Based Testing) の試験形式による国家試験等が医学教育において大きな課題となって

いる。

また、2020年の「医師国家試験改善検討部会」においても、医師国家試験へのコンピュータ制(CBT)の導入が示され、これらの実現に向けた取り組みが喫緊の課題となっている。今後、CBT 医師国家試験を実装するためには、CBT 試験システムの構築や実施のためのロジスティクス、動画・音声等を用いたCBT問題の作成体制、CBT 問題の評価・分析およびライブラリの構築などについて検討し、実際の CBT 医師国家試験を想定した検証をすることが必要である。

さらに、これまでの医師国家試験出題基準に関する検証や、外国の医学部を卒業した者の医師国家試験の申請数が増えつつあることも踏まえて、予備試験と日本語診療能力調査が課題として挙げられている。

そこで、各大学における診療参加型臨床実習の充実とそれにおける医学生の評価をより適正に実施すること、さらには動画・音声等を取り入れた実臨床に近い問題による CBT 化された医師国家試験により、「知識」だけでなく「技能」「態度」を含めた臨床能力を適正に評価することを一体的に考えることが重要である。

このように、各大学の医学教育における臨床実

習を充実させ、学生を適正に評価すること、さらに CBT 医師国家試験において知識・技能(一部)の評価が可能となることにより、国民から信頼される臨床能力に優れた医師を養成することができる。

本研究は、(1)CBT 医師国家試験の導入の実現に向けた課題の抽出、(2)技能や態度の評価、(3)医師国家試験出題基準の検証、(4)外国の医学部を卒業した者に対する予備試験と日本語診療能力調査の課題に関して研究を行い、国において医師国家試験と医学教育に関する検討が行われる際の資料として活用できる知見を得るとともに、今後の医師国家試験と医学教育の改善に向けて提言することを目的とする。

B. 研究方法

本研究は、「CBT 医師国家試験の導入の実現に向けた課題の抽出(研究Ⅰ:CBT 医師国家試験の導入の実現に向けた研究)」「診療現場における医学生の技能や態度の評価の方法におけるメリットとデメリットの抽出(研究Ⅱ:診療現場における医学生の技能や態度の評価に関する研究)」「令和6年度版医師国家試験出題基準改定に伴う第118回試験の検討、医師国家試験における「必修の基本的事項」「医学総論」等に関する検証を通じての

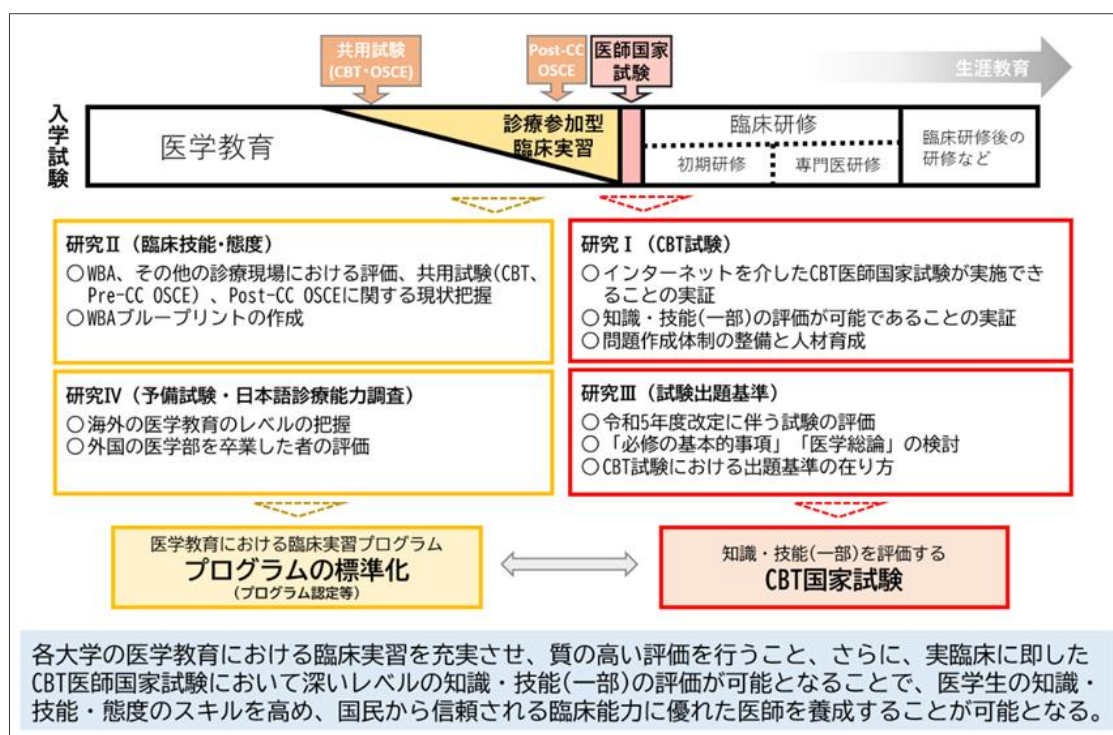


図1 河北研究全体構想

提言(研究Ⅲ:医師国家試験出題基準に関する研究)」「外国の医学部を卒業した者に対する予備試験および日本語診療能力調査について課題の整理と改善策の提言(研究Ⅳ:外国の医学部を卒業した者に対する予備試験等に関する研究)」の4つの課題に関して、以下の研究分担者のもと研究を行い、それぞれ知見を取りまとめた。

研究課題	分担者
I. CBT 医師国家試験の導入の実現に向けた課題の抽出	岡崎、松山、浅田、伴、久保
II. 診療現場における医学生の実能や態度の評価の方法におけるメリットとデメリットの抽出	錦織、小西、伴、松山、高村
III. 令和6年度版医師国家試験出題基準改定に伴う第118回試験の検討、医師国家試験における「必修の基本的事項」「医学総論」等に関する検証を通じての提言	小西、小松、高村、錦織、伴
IV. 外国の医学部を卒業した者に対する予備試験および日本語診療能力調査について課題の整理と改善策の提言	奈良、内藤、森、藤川

1. CBT 医師国家試験の導入の実現に向けた研究(研究Ⅰ)

CBT 医師国家試験の導入を実現させるためには、CBT 問題における動画・音声等の質に関する課題、CBT 試験システムの操作上の課題、サーバ構築やセキュリティ面における課題、試験会場や設備面の課題、不正防止対策等の運営面の課題、CBT 問題の作成・管理に関する課題などを整理して解決策を検討する必要があることから、先行研究に続き、全国の82大学医学部に呼びかけて、医学部の5年生または6年生を対象に、大学および大学以外の会場においてCBTトライアル試験を実施し、CBT 試験システムの課題、PC 環境や実施環境等に関して調査した。

試験終了後に受験者を対象としたアンケート調査を行った。また、関係の教職員に対してもアンケート調査を行った。

CBT 医師国家試験に向けて十分な量の問題数を確保するために、前年度に検討した「CBT 試験の一貫した運用を行うためのプロトタイプ」を継続的に検討し、CBT 問題の作問とライブラリシステムの構築を図った。

の構築を図った。

CBT 問題の作問については、生成 AI を活用した大量の派生問題の作成を行った。

また、プール問題のライブラリ化の構築に向けて、2025年度のトライアル試験問題については、項目反応理論(Item Response Theory : IRT)に基づく分析と等化を試みた。

2. 診療現場における医学生の実能や態度の評価に関する研究(研究Ⅱ)

日本の卒前医学教育における Workplace-Based Assessment(WBA)導入に関する国際ワークショップを開催し、WBA 導入の背景と目的、国内外の WBA 実施状況と課題等について議論した。

また、このワークショップで得られた論点整理を踏まえ、研究班内で原案を作成し検討を重ねて、臨床実習において、ある程度標準化された方法で観察評価を行う、Genba-Based Assessment (GenBA)と呼ぶ実施指針案を作成した。さらに、GenBA の制度設計と並行して、総括評価の導入にあたり重要な留意点を検討するため、障害者支援と医療安全の専門家にインタビューおよび討論を実施した。

3. 医師国家試験出題基準に関する研究(研究Ⅲ)

令和6年版医師国家試験出題基準の医学各論の大項目ⅠからⅩⅢまでの備考欄に記載された記載内容を3分類(追加病名、検査・術式・治療等、説明・付加症状)に整理した。各項目について、今後の出題基準改定時に整理対象となり得るかを探索的に検討し、「削除可能な項目」または「検討を要する項目」に分類した。

さらに、前年度に続けて、医学総論(150問)、必修の基本的事項(100問)の各設問について検証パネラー(13名)がレベル分類の検証を行った。レベル分類は、臨床実習前の学修で回答可能(レベルⅠ)、臨床実習中の学修成果を問う(レベルⅡ)、臨床研修で修得すべき(レベルⅢ)、専門研修で修得すべき(レベルⅣ)の4段階を設定した。

4. 外国の医学部を卒業した者に対する予備試験等に関する研究(研究Ⅳ)

前年度の調査結果で、外国の医学部出身者の医療活動に課題を指摘した研修病院が存在した。そこで、課題を指摘した 5 病院に対し、Web会議もしくは書面調査を行って課題を検討した。また、外国籍の医師に対する日本の一般市民の選好と懸念を明らかにするため、一般市民を対象とした Web 横断調査を実施し、回答を解析した。

C. 研究結果

1. CBT 医師国家試験の導入の実現に向けた研究(研究Ⅰ)

1) CBT トライアル試験の実施

CBT トライアル試験を全国の 60 大学と 1,647 名の医学生に参加により実施した。大学会場形式は 56 大学、大学以外の一般会場形式は 5 大学が参加した。

試験実施後のアンケートでは、医学生の 1,621 名からの回答が得られた(回収率 98.4%)。アンケートにおいて、マルチメディアを用いた問題が医師国家試験に有用であるかについては、「有用である・概ね有用である」が 87.0%であった。医師国家試験が CBT 化されてほしいかについては、「そう思う・ややそう思う」が 46.1%であった。CBT システムの操作については、「問題はなかった・概ね問題はなかった」が 96.1%であった。

CBT システムについては、CBT トライアル試験を通して、TAO を使用した CBT 形式での試験の実施は、概ね問題なく実施することができた。

2) CBT 問題の作問・ライブラリシステムの検討

前年度、Moodle を用いた作問システムについて検討し、本年度はこのシステムを基本とし、問題プールまでを含めたシステムとしての検討・構築を行った。

問題プールについては、Moodle をプラットフォームとし、医師国家試験出題基準の領域(医学各論・医学総論・必修)および大・中項目で区分したライブラリを設け、CBT システム(TAO)と Moodle を連携させた仕組み(MooTAO)を構築した。先行研究から作成してきた CBT 問題や今年

度の新作問題、生成 AI による派生問題をプールした。

3) CBT 問題の作問と評価

生成 AI による大量派生問題の作成に取り組み、カスタム GPT の初期型から、プロンプトを改良することにより、元問題と同じ症候や疾患をテーマとしつつ、患者の病歴、現症、検査所見、設問内容(病態、疫学、診断、治療、予後等)、選択肢に多様性を持つ派生問題が作成されるようになった。公衆衛生領域の一般問題は、派生問題ではなく、医師国家試験出題基準の記載用語をキーワードとした作問題の場合でも、概ね採用可能な問題として作成できた。マルチメディアのうち静止画については、元問題の本文中に含まれる臨床情報をカスタム GPT が推測することで、静止画そのものを直接読み取らなくても派生問題を作成できた。一方、動画・音声付問題については、元問題の本文中の臨床情報が限られていることに加え、動画や音声そのものをカスタム GPT が読み取ることができなかった。そこで、作問者が動画・音声の内容について 100~200 字程度の説明文を作成し、その説明文をカスタム GPT に読み取らせることで、派生問題として作成することができた。カスタム GPT を導入することで、昨年度までの作成ペースを上回る問題数を作成することができた(表1)。

表1 2021~2025 年度の累積作問数

年度	2021	2022	2023	2024	2025
作問数	200	200	200	367	1,790
累積問題数	200	400	600	967	2,757

CBT 問題の評価については、2025 年度の CBT トライアル試験の問題として出題した 200 問に対し、項目反応理論による分析を行い、A 問題(医学各論)、B 問題(必修)、C 問題(医学総論)の各問題の内容に応じた困難度・識別力の特徴が確認され、テスト情報関数からは想定される能力値の範囲で高い測定精度が保たれていることが示された。また、等化係数法による等化係数を算出した(表2)。

表2 等化係数の算出

	b の平均	b のSD	$\hat{R}$	$\hat{L}$
A 問題	-0.270 -0.563	1.324 1.179	1.124	0.363
B 問題	-0.674 -1.062	1.337 1.087	1.078	0.471
C 問題	-0.535 -0.753	1.407 1.232	1.142	0.324

2. 診療現場における医学生の技能や態度の評価に関する研究(研究Ⅱ)

日本国内では82校中65校がWBAを形成評価として実施しているが、総括評価としての利用はないこと、および米国ではStep 2 Clinical Skills 廃止後にmini-CEXやGlobal Practice Assessmentが総括評価に活用されていることなどが共有された。これらを踏まえ、全国の医学部で活用し得るWBAの雛形としてGenBAの実施指針案を作成した。GenBAでは、臨床実習における幅広い能力を評価対象候補としつつ、総括評価としてはコミュニケーションおよびプロフェッショナリズムを扱い、それ以外の臨床能力については各大学・各診療科の文脈に応じて形成評価として実施する構成とした。

GenBA 実施指針案の主な内容

- 全診療科・必修／選択・学外／海外実習を含む評価枠組みの提示
- 自己評価と指導医評価を組み合わせた運用案の提案
- 評価者支援として、不合格基準を共有する動画教材等の必要性の提示
- 「警告」→「再教育対象」→「不合格」という段階的対応案の整理
- 再教育の方法として、カウンセリング、再履修、特別再教育プログラムの例示

識者インタビューでは、障害者支援の観点から、医師として譲れない本質的能力と、合理的配慮により補う非本質的能力を区別する必要性が示

され、また医療安全の観点からは、患者安全にとって危険な学生の認定をコミュニケーションおよびプロフェッショナリズムのみで構成してよいかについては慎重な検討が必要であることが示された。

3. 医師国家試験出題基準に関する研究(研究Ⅲ)

令和6年版医師国家試験出題基準の医学各論における備考欄の検証については、備考欄の記載が存在する箇所は330か所であるが、1つの備考欄に複数の病名、検査・術式・治療等があるため延べ476項目であった。476項目のうち、「削除可能な項目」は344項目、「検討を要する項目」は132項目であった。

医学総論および必修の基本的事項の検証については、必修問題において、レベルⅠ問題の出題割合が50%以上であった項目は、「1.医師のプロフェッショナリズム」「3.診療情報と諸証明書」「5.人体の構造と機能」「10.臨床判断の基本」「17.生活習慣とリスク」の5項目であり、その全てが一般問題であった。医学総論においては、レベルⅠ問題の出題割合が50%以上であった項目は、「Ⅰ.保健医療論」「Ⅱ.予防と健康管理・増進」「Ⅲ.人体の正常機能と構造」の3項目であった。

4. 外国の医学部を卒業した者に対する予備試験等に関する研究(研究Ⅳ)

外国の医学部出身者の医療活動に課題を指摘した5病院に対する調査において、言語・疎通の問題、母国における臨床経験が尊重されないため他の職員との診療上の衝突によるストレス、臨床能力が低い、日本の医療制度や診療スタイルに不慣れ、臨床実習が十分確保されない等の課題が示された。

一般市民を対象としたWeb横断調査では、約半数から3分の2は日本人医師を選好しており、特に「コミュニケーション能力」65.5%、「共感性」62.9%、「患者の話をよく聞く姿勢」62.9%でその傾向が強かった一方で、「専門性」では47.8%が「国籍による選好なし」であった。

D. 考察

1. CBT 医師国家試験の導入の実現に向けた研究(研究Ⅰ)

1) CBT トライアル試験の実施

CBT トライアル試験は、先行研究を開始した 2021 年当初(総受験者数 321 名)と比較して、受験者・参加協力大学は増加しており、将来の CBT 医師国家試験の実装を見据え、実施方法や運用体制等について、より具体的な知見や経験を積み重ねた。CBT トライアル試験による継続した調査研究の意義もさらに増大するものと考えられる。トライアル試験に参加した受験者によるアンケート結果では、CBT 形式に変更することに伴い、試験対策や教材の不足を指摘する意見は複数あったが、CBT トライアル試験を受験したこと自体に対しては肯定的な意見が 8 割以上であったことから、学生の立場として一定の意義を感じていたと考えられる。そのため、医師国家試験を見直しするにあたっては、現在の医学部教育の実態を調査し現状を把握したうえで検討する必要がある。

また、CBT トライアル試験では、事前に CBT システムの動作に慣れるための教材も用意した、しかし、共用試験 CBT と同様に、事前に一定の時間をとってシステムの操作に慣れる機会を与えることが必要になると考えられる。多数の学生から「問題用紙に直接書き込めない」との意見があった。現状では、メモ用紙を配布してはいるものの、学生にとっての利便性は紙の問題用紙の方が高いことがうかがえる。

試験会場については、大学会場形式・一般会場形式を問わず、使用する教室や人数規模によって、座席配置や距離についての課題が残るとともに、試験における不正行為を防止するための仕組みは継続的な検討も必要であると考えられる。

タブレット端末を利用した CBT トライアル試験は、試験自体の運用については大きなトラブルもなく実施可能であることが確認できた。特に Wi-Fi の強度やバッテリーについて、特段の問題がなく運用可能であることが確認できたことは大きな意味がある。タブレット端末を利用した医師国家試験の実施の課題として、バッテリーの容量、イヤホンジャックの有無、タッチペンの反応、受験者の首や目の

負担などが考えられることから、その対応を検討する必要がある。

2) CBT 問題の作問・ライブラリシステムの検討

Moodle のデータベースモジュールを基盤として作成したことで、問題に対して付与する属性を比較的自由に追加・変更可能となっている。CBT 医師国家試験の運用面に関する検討と合わせ、根幹となる作問についての支援システムも継続的に整えていくことが求められる。

3) CBT 問題の作問と評価

一定期間内に問題を量産するためには、適切な基礎学習とプロンプトによる指示のもとで管理された生成 AI の支援による派生問題作成が有用であることが示された。

動画・音声問題については、カスタム GPT による直接的な読み取りは困難であり、派生問題を適切に作成できない状況であった。これは CBT における動画・音声付き臨床問題が、生成 AI 時代におけるヒト臨床医の実践能力を評価するうえで、妥当な形式である可能性を示唆するものである。

IRT による分析結果から、A 問題と C 問題は、易しい項目から難しい項目まで幅広い困難度の項目を含み、おおよそ -2.0 から 1.0 までの能力値をもつ受験者をよく識別するテストであった。B 問題は、相対的に易しい項目が多いことが明らかとなった。また、2023 年度の等化後の項目母数を利用することで、2023 年度の受験者に対して 2025 年度と同一の尺度上で能力母数を算出することも可能となる。

2. 診療現場における医学生の技能や態度の評価に関する研究(研究Ⅱ)

日本における WBA の制度設計に必要な論点を具体化することができた。一方で、GenBA はまだ開発途中にあり、今後も継続してブラッシュアップしていく必要がある。とくに、警告・不合格基準の標準化、評価者間のばらつきへの対応、質保証の方法、再教育プログラムの運用、障害学生支援との連携、患者安全の観点が必要な課題として残されている。

3. 医師国家試験出題基準に関する研究(研究Ⅲ)

医学各論の備考欄については、延べ 476 項目の記載があり、とくに追加病名は 283 項目と最も多く、備考欄が単なる補足説明の欄にとどまらず、出題対象となり得る病名を追加的に記載する機能を有している可能性を示している。延べ 476 項目のうち 344 項目が削除可能な項目、132 項目が検討を要する項目と判断されたが、医学各論の備考欄の相当数が今後整理できる可能性を示している。

備考欄に記載された追加病名の中には、臨床研修開始前に求められる知識として重要であるものも含まれていたが、このような項目は、備考欄に残すか削除するかではなく、小項目本文に移動し、必要に応じて a,b,c 分類の対象として整理することが望ましい。備考欄の記載内容や位置づけを曖昧なままにしておくことは、出題基準の明確性やレベル分類との整合性の観点から課題となる可能性がある。

医学総論および必修の基本的事項については、今回の検討結果を踏まえ、保健医療や人体の構造と機能、診療情報の取扱いの基本等は、ブループリントにおける出題割合を減らし、その分を臨床実習開始前に実施される共用試験 CBT で一定程度出題するなどの調整が望ましいと考えられる。今回の検証データを毎年開催される出題委員等の会議で事前情報・知見として提供することで、問題作成時およびその後のブラッシュアップ時に、レベルⅠ問題作成数の減少に寄与できる可能性もあると考える。

4. 外国の医学部を卒業した者に対する予備試験等に関する研究(研究Ⅳ)

外国の医学部出身者の医療活動に課題を指摘した 5 病院に対する調査では、個々の研修医の背景、臨床経験、日本の医療文化への適応状況、受け入れ側の支援体制との相互作用の中で生じている可能性が示された。とくに重要なのは言語面である。実臨床で求められる日本語運用能力は、より高次の文脈依存的能力が求められる。そのため、予備試験、日本語診療能力調査に合格していても、実臨床では十分な疎通が困難な

場合があり得る。

一般市民を対象としたアンケートでは、一般市民にとっても、医師の専門性そのものに加えて、十分に意思疎通ができるかどうかは医療への信頼を左右する重要な要素であることが示された。診療現場における実践的日本語能力を継続的に評価し、必要に応じて補完的支援を行うことが重要である。

E. 結論

1. CBT 医師国家試験の導入の実現に向けた研究(研究Ⅰ)

試験環境については、大学会場形式と一般会場形式、それぞれの運用を確認できたことと合わせ、タブレット端末を利用することでの利点・課題点を整理した。

CBT問題の作問・ライブラリシステムについては、問題の整理・検索の機能を強化するための検討を行った結果、医師国家試験出題基準等に準じたプール問題の整理・検索が可能となった。作問は医師国家試験の運用に際しても要となる要素の一つである。試験運用の仕組み整備と合わせ、良質な問題を確実に整理するためのシステムについても継続的な検討が必要である。

CBT 問題の体系的な管理基盤を MooTAO 上で構築するとともに、カスタム GPT により臨床問題の派生問題等を短期間に多数作成できることを実証することができた。しかしながら、採用可能な問題は約 7 割であり、専門医による採否判定と校正は不可欠であった。また、動画・音声付き問題では生成 AI のみでの作問に限界があり、作業工程にヒトが介することを前提とした取り組みが必要と考える。一方、動画・音声が付与する視覚・聴覚情報の認識は、生成 AI 時代においてもヒト臨床医が引き続き育むべき臨床実践力ともいえる。CBT 導入に際しては、動画・音声付き臨床問題の積極的採用を検討すべきである。

IRT に基づくテスト運用では、等化の手続きを経た項目の集まりである項目プールをあらかじめ用意する必要がある。そこで本研究では、等化係数法による項目母数の等化の一例を示した。医師国家試験においても、異なる受験者集団が、異なる

る項目から構成される問題を、異なる日時に、異なる場所で受験したとしても、統一の枠組みで評価可能となるような IRT に基づくテスト運用は望ましい。なお、項目プールの安定的な維持と効率的な拡大のためには、項目内容の非公開が前提条件となる。

2. 診療現場における医学生の技能や態度の評価に関する研究(研究Ⅱ)

国際ワークショップを通じて、日本における総括的 Workplace-Based Assessment(WBA) 導入に関する主要論点を整理し、その知見を踏まえて、我が国の臨床実習における総括的 WBA の実施指針案として Genba Based Assessment(GenBA)を作成した。さらに、障害者支援および医療安全の専門家との意見交換を通じて、本質的能力の定義、合理的配慮、患者安全を踏まえた評価基準設定など、制度設計上の重要課題を明らかにした。

我が国の医学教育における総括的 WBA の導入に関する提言として以下のとおり取りまとめた。

提 言

医師資格の認定において、知識のみならず技能・態度を適切に評価することは重要な課題である。これまで、OSCE は標準化された技能評価手法として卒前医学教育における質保証に一定の役割を果たしてきたが、その実施には模擬患者、評価者、会場設営、運営体制など、多大な人的・物的資源を要することから、持続可能性および費用対効果の観点で課題がある。

一方、Workplace-Based Assessment(WBA)は、実際の診療現場において学習者を継続的に観察し評価することが可能であり、単発の試験では十分に把握しにくい技能や態度を、より実践に即した形で評価し得る手法である。本研究における全国 82 医学部の自己点検評価報告書の分析では、4 分の 3 を超える大学がすでに何らかの形で観察評価を実施していることが確認されており、WBA は一定の実装基盤を有していると考えられる。

もちろん、WBA を総括評価として活用する

ためには、評価者間のばらつきへの対応、評価基準の標準化、質保証、合否判定の妥当性、再教育との接続など、なお検討すべき課題が存在する。しかし、こうした課題は、WBA の制度化を妨げる理由というよりも、むしろ制度設計を進める中で解決していくべき課題である。

以上を踏まえ、今後、国家試験あるいは医師資格認定の過程における技能・態度評価のあり方として、WBA を適切に構造化・標準化し、プロフェッショナリズムやコミュニケーション能力の評価を中心に総括評価として実装することを積極的に検討すべきである。その際、OSCE で評価することが適切な標準化技能評価との役割分担を明確にしつつ、継続的かつ実践的な評価を可能とする持続可能な評価システムとして発展させていくことが望まれる。

3. 医師国家試験出題基準に関する研究(研究Ⅲ)

医学各論の備考欄に関する探索的検討により、令和 6 年版医師国家試験出題基準の医学各論に、延べ 476 項目の備考欄への記載が認められ、その内訳は、追加病名、検査・術式・治療等、説明・付加症状であり、追加病名が最も多かった。また、研究者による探索的検討では、延べ 476 項目のうち 344 項目が削除可能な項目、132 項目が検討を要する項目と判断された。また、第 116 回から第 118 回までの医学各論問題では、備考欄からの出題がそれぞれ 58 問、71 問、71 問認められ、備考欄記載が実際の出題にも関係している可能性が示された。本検討は研究者 1 名による予備的・探索的判断であり、備考欄記載の削除や改訂を直接提案するものではないが、今後、医師国家試験委員経験者、医学教育関係者、専門領域の専門家等による妥当性検証を行い、備考欄の役割を明確化したうえで、出題基準改定時の整理方針を検討する必要がある。

医学総論および必修の基本的事項については、必修の基本的事項の出題領域(18 の大項目)別で検証すると、レベル I 問題の出題割合が 50% 以上であった項目は、「1.医師のプロフェッショナリズム」「3.診療情報と諸証明書」「5.人体の構造と機能」「10.臨床判断の基本」「17.生活習慣とリスク」

の 5 項目であり、その全てが一般問題であった。次に、医学総論の出題領域(9つの章)別で検証すると、レベルⅠ問題の出題割合が50%以上であった項目は、「Ⅰ.保健医療論」「Ⅱ.予防と健康管理・増進」「Ⅲ.人体の正常機能と構造」、の3項目であった。これらについて、次回の医師国家試験出題基準改訂検討部会時にブループリントにおける出題割合を減らし、その分を臨床実習開始前に実施される共用試験 CBT で一定程度出題するなどの調整が望ましいと考えられる。また、毎年開催される出題委員等の会議で事前情報・知見として提供することで、問題作成時およびその後のブラッシュアップ時に、レベルⅠ問題作成数の減少に寄与できる可能性も期待される。

4. 外国の医学部を卒業した者に対する予備試験等に関する研究(研究Ⅳ)

外国の医学部卒業生の受け入れ、勤務状況に関しては大きな問題は見られなかった。一方、少数ではあるが外国の医学部出身医師の勤務状況に問題が認められたこと、および今後も外国の医学部出身医師の増加が見込まれることから、さらに詳細な調査研究を行い、より適正な外国の医学部出身医師の受け入れ制度を構築し、国民から信頼を得る必要があると考えられる。

今後、国において医師国家試験と医学教育について審議が行われる場合には、本研究で得られた成果を踏まえて議論されることを望む。さらに、国の施策として医師国家試験の CBT 化が決定された場合は、本研究で培った知見やノウハウが活用されることを期待する。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

なし

2. 学会発表

- 1) 2025 IEdTech Learning Impact Conference 講演: Trial for the Computer-Based Testing for the National Medical Licensing

Examination with Video and Audio-Enhanced Questions

- 2) 第 57 回日本医学教育学会大会, シンポジウム: 医師国家試験 CBT 化に向けた研究成果と今後の展望「CBT 形式での医師国家試験を実現するにあたっての試験システム検討」
- 3) 第 57 回日本医学教育学会大会, 一般演題: 医師国家試験 CBT 化の実装に向けた課題と考察～CBT トライアル試験の受験者アンケートの結果から～
- 4) 第 57 回日本医学教育学会大会, ポスター演題: 医師国家試験 CBT トライアルの IRT による分析—マルチメディア使用問題の結果を中心に—

H. 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

特記すべき事項なし

I. 研究協力者一覧(50 音順・敬称略)

2026年3月31日時点

<研究 1: CBT 問題作成・検討>

伊藤 真人	(自治医科大学 耳鼻咽喉科)
及川 沙耶佳	(秋田大学 大学院医学系研究科 先進デジタル医学・医療教育学講座)
大久保 由美子	(帝京大学 医学部 医学教育学講座)
海津 聖彦	(日本医科大学 思春期・小児医学／付属病院小児科／医学教育センター／臨床研修センター)
加賀 重亜喜	(山梨大学医学部附属病院 臨床教育部臨床実習センター、シミュレーションセンター)
角 総一郎	(自治医科大学 皮膚科学講座)
兼田 裕司	(自治医科大学 メディカルシミュレーションセンター 消化器一般移植外科)
菅野 武	(自治医科大学 医学教育センター医療人キャリア教育開発部門)
木村 友和	(名古屋大学医学部附属病院 泌尿器科)
金城 紀与史	(琉球大学 医学部 医学教育企画室)
黄 世捷	(聖マリアンナ医科大学 医学教育文化部門)
小佐見 光樹	(自治医科大学 地域医療学センター公衆衛生学部門)

後藤 理英子 (東京医科大学 医学・看護学教育推進センター)
 才津 旭弘 (自治医科大学 医学教育センター 医療人キャリア教育開発部門)
 相良 由佳 (医療法人社団 幸野メディカルクリニック)
 笹原 鉄平 (自治医科大学附属病院 感染制御部)
 澤幡 美千瑠 (自治医科大学 呼吸器内科)
 鋪野 紀好 (千葉大学大学院医学研究院 地域医療教育学)
 清水 郁夫 (千葉大学医学部附属病院 医療安全管理部)
 高橋 宏典 (自治医科大学 産科婦人科)
 田鎖 愛理 (岩手医科大学 衛生学公衆衛生学講座/全学教育推進機構教学 IR センター)
 蓮沼 直子 (広島大学 医学部附属医学教育センター)
 林 幹雄 (関西医科大学 教育センター)
 堀切 海都 (大間病院 内科・総合診療科)
 町田 幹 (日本医科大学付属病院 放射線科)
 松原 大輔 (自治医科大学 地域医療学部門)
 松本 卓子 (東京女子医科大学 統合教育学修センター基礎教育学)
 三原 弘 (札幌医科大学 医療人育成センター)
 宮道 亮輔 (自治医科大学 メディカルシミュレーションセンター)
 山田 和徳 (金沢医科大学 医学教育学)
 山本 祐 (自治医科大学 地域医療学センター総合診療部門)
 早稻田 勝久 (愛知医科大学 医学教育センター)
 渡辺 茅里 (自治医科大学 眼科学講座)
 渡会 瑠斗 (山形県立中央病院)

<研究2:WBA 調査・作成協力>

木村 武司 (名古屋大学医学系研究科)
 近藤 猛 (名古屋大学医学部附属病院 総合診療科/卒後臨床研修・キャリア形成支援センター)

<研究3:医師国家試験検証パネラー>

大久保 由美子 (帝京大学 医学部 医学教育学講座)
 金城 紀与史 (琉球大学 医学部 医学教育企画室)
 近藤 猛 (名古屋大学医学部附属病院 総合診療科/卒後臨床研修・キャリア形成支援センター)
 高橋 誠 (北海道大学 大学院医学研究院 医学教育・国際交流推進センター)
 高見 秀樹 (名古屋大学医学部附属病院 卒後臨床研修キャリア形成支援センター)
 永田 康浩 (長崎大学 生命医科学域(医学系))
 西屋 克己 (関西医科大学 教育センター)

野村 理 (岐阜大学 医学教育開発センター)
 蓮沼 直子 (広島大学 医学部附属医学教育センター)
 藤田 博一 (高知大学 医学部附属医学教育創造センター)
 松島 加代子 (長崎大学病院 医療教育開発センター)
 松山 泰 (自治医科大学 医学教育センター)
 三井 良之 (近畿大学 医学部教育センター)
 山本 憲 (順天堂大学 健康データサイエンス学部)

<事務局>

上田 茂 (公益財団法人日本医療機能評価機構)
 中田 祐太 (公益財団法人日本医療機能評価機構)

【参考文献】

1. 厚生労働省. 医師国家試験出題基準令和 6 年版
2. 河北博文(研究代表者). 厚生労働省科学研究費補助金事業「ICT を利用した医学教育コンテンツの開発と活用に向けた研究」令和 3 年度～5 年度総合研究報告書
3. 文部科学省. 医学教育モデル・コア・カリキュラム

