

CBT 医師国家試験の導入の実現に向けた研究

研究分担者 岡崎 仁昭 (自治医科大学医学教育センター 顧問)
松山 泰 (自治医科大学医学教育センター 教授)
浅田 義和 (自治医科大学医学教育センター 准教授)
久保 沙織 (東北大学教育学研究科 准教授)
伴 信太郎 (愛知医科大学医学教育センター 特命教育教授)

研究要旨

2020 年の「医師国家試験改善検討部会」において、医師国家試験へのコンピュータ制(CBT)の導入が示され、2021 年度～2023 年度 厚生労働省科学研究「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」(以下、先行研究)において、画像や音声、動画等を用いた CBT 問題によるインターネットを介した CBT 試験の試行が行われ、概ね円滑に実施できることを実証するとともに、CBT 医師国家試験の実装に向けて様々な課題について検討を行った。

本研究では、これまでの先行研究での実績を踏まえ、将来の CBT 医師国家試験を早期に実装するため、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より具体的に検討し、実際の CBT 医師国家試験を想定して検証することを目的に、先行研究での知見や各課題について解決策を検討し改善するとともに、全国 82 医学部を対象にした CBT トライアル試験の実施を通して、CBT 医師国家試験の実現に向けた課題と改善策を整理した。

CBT トライアル試験では、全国 52 大学(1,248 名)の参加により実施した。このうち、現行の医師国家試験と同様に外部会場方式の実施を想定して、日本医療機能評価機構の多目的ホールで 1 回実施した。試験問題は、現行の医師国家試験出題基準に基づき、より実臨床に近い問題となるよう、動画・音声等を取り入れた 200 問を作成し、インターネットを介した CBT 試験システム(TAO)を利用した。本年度は、各受験者が同時に複数の大学から CBT システムにアクセスすることを可能にしたことで、同日に 2 大学での実施が 9 日程、同日に 3 大学での実施が 2 日程と、複数の大学において同日に実施した。トライアル試験は大きなトラブルもなく概ね円滑に実施することができた。

また、CBT 問題作成体制の構築、CBT トライアル問題作成マニュアルの作成、問題の回答方法や不正防止対策、試験監督のルール等を整理し、マニュアルを作成するなど、医師国家試験を安全に効率的に運営するため、具体的な方策を検討し、さらに将来の CBT 医師国家試験での運用を見据え、CBT 問題の評価・分析とプール化、実施形式、実施費用の試算、CBT 試験を効率的に実装できるデータベースおよび試験運営システムの構築等について総合的に検討した。将来の CBT 医師国家試験を円滑に実施するための知見を整理するとともに、CBT 医師国家試験が実装された際に、活用可能なシステムおよびマニュアルであることを示すことができた。

さらに、我が国の国家資格試験の CBT 化について、既に CBT 形式が導入されている、または導入が予定されている国家資格試験等の情報収集を行い、医師国家試験の CBT 化の実現に向けて参考にした。

本研究を通して、CBT 医師国家試験の実装に向けて、医師国家試験に関連する検討部会における資料として活用できる知見を得ることができた。

A. 研究目的

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(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。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て検証することが必要である。

2021年度～2023年度 厚生労働省科学研究「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(研究代表者:河北博文)」(以下、先行研究)において、画像や音声、動画等を用いたマルチメディア形式の試験問題を作成して、インターネットを介した CBT トライアル試験を実施し、46大学で円滑に実施できることを実証したが、医師国家試験にCBT方式を本格的に導入するためには、CBT 問題における動画・音声等の質に関する課題や、CBT 試験システムの操作上の課題、サーバ構築、セキュリティ面における課題、試験会場および設備面の課題、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様々な課題が上がった。

本研究では、CBT 医師国家試験の導入を実現するために、先行研究での知見や各課題について解決策を検討し改善するとともに、全国82医学部を対象にした 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、さらに具体的に課題を抽出し改善策を検討することで、医師国家試験に関連する検討部会における資料として活用できる知見を得るとともに、ICT を利用した医師国家試験の実現に向けて提言することを目的とする。

B. 研究方法

1. CBTトライアル試験の実施

1) CBT 問題・TAO 設定と大学への案内

(1)概要

CBT の実施システムには、先行研究と同様に TAO(OAT 社・ルクセンブルク)を用いた。TAO 以外のシステムについては、Moodle 等の LMS を利用する案なども再度確認したが、将来的なメンテナンス性、CBT 専用のシステムとしての信頼性などを踏まえて、TAO を継続的に利用することとした。

(2)CBT の全体設計

CBT の問題は先行研究と同様、A 問題(各論)75 問、B 問題(必修)50 問、C 問題(総論)75 問の合計 200 問として準備を行った。これを用いて、CBT の準備設定を行った。

1)大学での実施

本年度の研究においては、基本的に先行研究と同様、各大学の PC 教室等を利用して、試験の運営を当該大学の職員等に協力いただき、受験者は当該大学の学生として参加を募り、大学ごとに実施した。

CBT トライアル試験の実施にあたっては、先行研究での参加した大学 46 大学に加え、全国のすべての 82 大学医学部に対して参加実施を依頼した。

2)会場形式での実施

現行の医師国家試験は、2 日間の試験を、大学の教室等の広い会場を借用して実施している。これは前述のような各大学の PC 教室等をそのまま利用して実施することとは異なり、大学入学共通テストのように他大学や他施設の会場を借用して、試験会場を構築する形式となる。本年度は、この会場形式についても費用の試算と合わせて運用のトライアルを実施した。

先行研究では、3 年間の CBT トライアルを実施するにあたり、会場および試験当日の運営をすべて大学に協力依頼することとしていた。しかし、本研究を行うにあたり、一定数の大学では PC の設置やネットワーク環境などを十分に整えることがで

きず、参加を見送ったという事情もあったことから、本年度は、大学以外の会場（多目的ホール）を借用し、試験運営を本研究の事務局が行う、会場形式を検討した。さらに、実施にあたって、会場周辺の複数大学の医学生を中心に呼びかけた。

会場形式で実施するにあたり、TAO の機能の一つであるプロクター機能を用いた試験の運営管理を試みた。プロクター機能を利用することで、試験の不正防止、画面ロックの機能などをより詳細に対応可能となる。一方、この機能を利用する場合はプロクターとして試験監督が手動で開始・終了の処理を行うことが必須となる。運用に際しての人的な負荷が増加する欠点はあるが、実際の医師国家試験本番での運用も見据え、会場形式で実施する 1 回に限り、プロクター機能を試験的に活用して、トライアル試験を実施した。

3) CBT トライアル試験の評価

(1) 学生アンケートの実施

CBT トライアル試験の受験者に対して、先行研究と同様に、CBT 試験システムの使用感、動画・音声等を取り入れた CBT 問題の感想、医師国家試験の CBT 化への期待など、C 問題の回答後にアンケートを実施した。設問は以下の通りである。

- Q1. 学年を解答してください。
- Q2. CBT 試験にて操作は問題なく行うことができましたか。
- Q3. 動画ファイルの再生はスムーズにできましたか。
- Q4. 音声ファイルの再生はスムーズにできましたか。
- Q5. 試験問題の文章の読みやすさはどうでしたか。
- Q6. 試験時間はどうでしたか。
- Q7. 試験問題の難易度はどうでしたか。
- Q8. CBT 試験を受験して良かったですか。
- Q9. 今回の試験問題には、動画や音声などのマルチメディアファイルを利用した問題を取り入れました。このような問題は、医師国家試験に有用であると思いますか。
- Q10. 今後、医師国家試験の CBT 化に期待しますか。
- Q11. 以下の機能で役に立ったものをお答えください（複数選択可）。
- Q12. 配布されたメモ用紙・筆記用具は役に立ちましたか。
- Q13. CBT試験を受験して、全体を通じて、ご意見やご要望があればお聞かせください。

(2) 教職員

研究協力大学の担当者に対して、試験実施後のアンケートを依頼した。設問は以下の通りである。

- Q1. CBT トライアル試験の準備や運営はスムーズにできましたか。
- Q2. CBT トライアル試験を貴大学で実施するにあたり、準備や運営に際し苦勞されたことがあれば教えてください。
- Q3. 今回の試験問題には、動画や音声などのマルチメディアファイルを利用した問題を取り入れました。このような問題は、医師国家試験に有用であると思いますか。
- Q4. 今後、医師国家試験の CBT 化に期待しますか。
- Q5. CBT トライアル試験を実施して良かったですか。
- Q6. CBT トライアル試験を実施して、参加学生の受験様子などから何か気になったこと、感じたことはありますか。
- Q7. CBT トライアル試験の実施を通じて、ご意見やご要望があればお聞かせください。

4) CBT 問題の評価

2024 年度の CBT トライアルで出題された 200 問について、A 問題(75 問)、B 問題(50 問)、C 問題(75 問)のそれぞれで、古典的テスト理論に基づく項目分析、および IRT による項目母数の推定を実行した。まず、各項目の要約統計量を確認した上で、古典的テスト理論に基づき項目困難度（通過率）と項目識別力（item-total correlation: IT 相関）を求めた。その後、IRT の 2 母数ロジスティックモデル(two parameter logistic model: 2PLM)を適用して困難度母数と識別力母数を推定した。なお、IRT の分析においては、安定的な推定のために、IT 相関が 0.2 以上の項目のみを用いた。

2. CBT 問題の作問と体制

1) CBT 問題の作問

現行の医師国家試験は、「必修の基本的事項」「医学総論」「医学各論」の 3 領域で構成されており、各日 200 問、2 日間で計 400 問が出題される。出題基準においては、各領域について大・中・小項目でテーマを明示し、さらにブループリント(医師国家試験設計表)により各章および大項目の出

題割合が規定されている。

本研究における CBT トライアルは出題割合を順守しながら、参加のしやすさを考慮して 1 日 200 問で完遂できる試験とした。次節に記載する手法により研究協力者をリクルートし、協力者の専門領域を踏まえて、各章・大項目の出題基準に則して問題を割り当てた。

また、マルチメディア素材(動画・音声等)を活用した問題については、作問者が素材を使用できる場合に自主的に作成する形式とし、出題領域や問題数を事前に限定しなかった。問題は Moodle の提出フォーム上で作成され、写真・動画・音声等の素材については、データ容量が小さいものは Moodle に直接添付、大容量のものはクラウド共有システムを用いて提出された。

2) 作問者およびブラッシュアップメンバーの選出

CBT 問題の作問者は研究分担者(岡崎、松山)および研究協力者とした。研究協力者は、先行研究において、マルチメディアを使用した医学教育コンテンツ(臨床問題教材)を作成したメンバー、および同研究における人材育成ワークショップ(第 54 回・55 回日本医学教育学会大会、第 83 回・86 回医学教育セミナーとワークショップ)の参加者、研究分担者の推薦による人物から選出した。さらに、作問者とは別に、作問した CBT 問題を客観的に評価し、問題のブラッシュアップを主に担当する研究協力者を別途選出した。

いずれの協力者も、本研究の目的と趣旨に同意の上で参加した。リクルート期間は 2024 年 4 月から 7 月に設定した。

3) CBT 問題作成マニュアルの作成

先行研究の医学教育コンテンツ教材作成マニュアルを基盤とし、令和 6 年度版医師国家試験出題基準に新たに記載された医学総論の小項目ごとの適切なタクソノミーレベル 1 を参照し、さらに記述形式の統一を目的に、用語、表記、単位、基準値等を医師国家試験の公開問題に準じて整理し、CBT 問題作成マニュアルとして再編した。また、マルチメディア素材の記録・保存・編集方法、および作問への活用法についても、先行研究の教材作成マニュアルを踏まえて、CBT 用にブラッシュアップした

内容を記載した。

3. 国内の国家資格試験における CBT 試験導入の状況

我が国では、コロナ過での感染対策における様々な行動制限等により、会議やイベント等での ICT の活用状況が大きく変わり、国内の民間企業が実施する資格試験において、CBT 形式で実施する資格が増加している。このような背景から、我が国の医師国家試験以外の他の国家資格試験において CBT 形式を導入している現状を把握し、医師国家試験の CBT 導入に向けて参考にするため、CBT を導入している国家資格試験の管轄である省庁の担当者を訪問し、ヒアリングを行った。

なお、厚生労働科学研究の「保健師助産師看護師国家試験におけるコンピュータの活用に関する研究(22IA2008)」の分担研究「保健師助産師看護師国家試験におけるコンピュータの活用に関する研究」の成果をはじめ、Web で公開されている情報も参考とした。

C. 研究結果

1. CBT トライアル試験の実施状況

1) 実施結果

(1) TAO の設定および大学への案内

TAO は定期的にバージョンアップが行われており、先行研究の研究期間中にユーザーインターフェースの刷新が行われ、特に受験者の操作画面が大きく変更となった。なお、以降、トライアル試験に参加する者を「受験者」と呼ぶ。先行研究では研究期間の途中だったこともあり、3 年間を通じて旧バージョンの操作画面を使い続けることとしていた。今後は新バージョンが利用されていくこと、直感的な操作として分かりやすくなっていることなどを踏まえ、新バージョンの TAO でトライアルを開始した。

(2) CBT の全体設計

① CBT の問題設定

設問は紙媒体の試験と同様に文字・画像を用いた問題と、動画や音声を含めて情報を提示した問題の 2 通りを用いた。

動画・音声を含む問題は A 問題で 7 問、B 問題で 10 問、C 問題で 11 問の合計 28 問を出題した

(全体の 14%)。なお、画像を含む問題は A 問題で 48 問、B 問題で 19 問、C 問題で 21 問の合計 88 問であった(44%)(表1)。

表 1:マルチメディア使用問題

	マルチメディア	問題数	割合
A 問題 (75 問)	動画・音声	7	9.3%
	画像のみ	48	64.0%
B 問題 (50 問)	動画・音声	10	20.0%
	画像のみ	19	38.0%
C 問題 (75 問)	動画・音声	11	14.7%
	画像のみ	21	28.0%
全体 (200 問)	動画・音声	28	14.0%
	画像のみ	88	44.0%

先行研究での実施方法と同じく、選択肢の表示順番は医師国家試験の選択肢表示順番に揃えることとした。このため、TAO 自体には選択肢の順番をランダムに入れ替える機能が存在してはいるものの、今回はその利用を見送った。

同じく、先行研究での実施と同様、出題の順番は受験者ごとにランダムになるように設定した。なお、A 問題、B 問題、C 問題それぞれの区分において、一般問題と臨床問題とが出題されるため、先行研究での知見も踏まえ、問題のシャッフルはそれぞれの問題区分の中でのみ実施した。例えば、一般問題 30 問と臨床問題 20 問であれば、最初の 30 問の中で一度シャッフルし、その後続く 20 問の中でシャッフルを行う設計とした。

また、先行研究からの変更点として、TAO が新バージョンに移行したことで、連問を扱うことが少し困難となったため、連問は共用試験 CBT にもあるような、「次の問に進むことで、追加情報が表示される」形式とした。追加情報が提示されることで、ほぼ自動的に 1 つ前の設問の正解が判明するため、この形式では後戻りをできなくする必要がある。新バージョンの TAO においても機能としては設定可能であるが、テスト内の一部の問題のみに適応することができないため、本年度のトライアル試験では連問形式にしない、あるいは連問となる場合も前問の答えに直結する情報を提示しない作りとした。

② CBT の画面設定

TAO が新バージョンとなり、受験者に対して回答時に表示される画面のレイアウトも大きく変更されることとなった。これに際し、問題や画像・動画の表示形式を可能な限り統一し、可視性を高めることを検討した。具体的には、画像や動画を含まない問題は 1 カラムですべて表示し、画像やレイアウトを含む問題は 2 カラムとして左側に本文と選択肢、右側に画像や動画を配置するようにした(図1)。



図1 画像や動画を含む問題のレイアウト例

③ CBT の動作設定・動作検証

先行研究と異なり、TAO が新バージョンとなったことで、新たに公開セッションを扱うことができるようになった。これはユーザとしてのログインが不要であり、誰でも自由に利用可能なテストを作る機能となる。先行研究では、各参加予定大学に動作検証用のアカウントを配布し、試験開始前までの接続確認、受験者への操作練習・説明に利用可能な状態を用意していた。このため、各大学用のアカウント発行や連絡などの手間が生じていたが、今回は、公開セッションを利用し、事前に URL を発行・配布することで、申込みの時点で動作検証用のテストにアクセスすることを可能とした。

以上を踏まえて、TAO の設定を行った。出題する問題については、研究分担者および事務局による手作業にて TAO システムへ入力を行った。なお、受験者の画面と異なり、問題作成に関する画面は基本的に大きな変更はない状態であったため、先行研究での経験を踏まえ、研究分担者 1 名ではなく、複数人で作業を分担することができた。

さらに、問題作成についても、複数の問題作成者により数多くの CBT 問題を作成するため、問題

作成する方法について検討を行ったが、CBT問題の作成については TAO 以外のシステムで行い、作問者の利便性を高めることとした。

受験者の登録やグループ分けは新規のインターフェースに切り替わったが、実行可能な内容については大きな変更点もなく、基本的には前回研究班の手順を踏襲することで実施が可能であった。

なお、細かな操作手順については管理者にのみ影響する内容であり、本稿では詳細については省略する。

(2) 全体の実施結果

CBTトライアル試験は 52 大学からの受験があり、合計 1,248 名の受験者となった。A 問題(75 問)、B 問題(50 問)、C 問題(75 問)および全体(200 問)の平均点は、A 問題 33.7 点、B 問題 29.9 点、C 問題 36.2 点、全体で 104.0 点であった。点数に関する概略および学年ごと受験者数と平均点は表3、表4で示す。受験者数は 50 名以上の参加大学が 9 大学あった一方、21 大学は参加者数が一桁であった。本年度は同一受験日の制限が緩和されたこともあり、同日に 2 大学での実施が 9 日程、同日に 3 大学での実施が 2 日程と、複数の大学において同日に実施し、円滑に行われた(表2、表3、表4)。

20	帝京大学	東京都	私立
21	東京医科大学	東京都	私立
22	東京女子医科大学	東京都	私立
23	日本大学	東京都	私立
24	日本医科大学	東京都	私立
25	横浜市立大学	神奈川県	公立
26	聖マリアンナ医科大学	神奈川県	私立
27	東海大学	神奈川県	私立
28	新潟大学	新潟県	国立
29	金沢医科大学	石川県	私立
30	山梨大学	山梨県	国立
31	浜松医科大学	静岡県	国立
32	愛知医科大学	愛知県	私立
33	三重大学	三重県	国立
34	滋賀医科大学	滋賀県	国立
35	京都大学	京都府	国立
36	京都府立医科大学	京都府	公立
37	大阪医科薬科大学	大阪府	私立
38	神戸大学	兵庫県	国立
39	兵庫医科大学	兵庫県	私立
40	和歌山県立医科大学	和歌山県	公立
41	川崎医科大学	岡山県	私立
42	広島大学	広島県	国立
43	徳島大学	徳島県	国立
44	高知大学	高知県	国立
45	九州大学	福岡県	国立
46	福岡大学	福岡県	私立
47	佐賀大学	佐賀県	国立
48	長崎大学	長崎県	国立
49	熊本大学	熊本県	国立
50	大分大学	大分県	国立
51	宮崎大学	宮崎県	国立
52	鹿児島大学	鹿児島県	国立

表2 2024 年度 CBTトライアル試験参加大学

	大学	都道府県	国立/私立
1	北海道大学	北海道	国立
2	旭川医科大学	北海道	国立
3	札幌医科大学	北海道	公立
4	弘前大学	青森県	国立
5	岩手医科大学	岩手県	私立
6	東北大学	宮城県	国立
7	東北医科薬科大学	宮城県	私立
8	秋田大学	秋田県	国立
9	山形大学	山形県	国立
10	福島県立医科大学	福島県	公立
11	自治医科大学	栃木県	私立
12	獨協医科大学	栃木県	私立
13	群馬大学	群馬県	国立
14	埼玉医科大学	埼玉県	私立
15	千葉大学	千葉県	国立
16	国際医療福祉大学	千葉県	私立
17	東京医科歯科大学	東京都	国立
18	順天堂大学	東京都	私立
19	昭和大学	東京都	私立

表3 各テストの平均点・標準偏差など

	A 問題	B 問題	C 問題	全体
平均点	37.7	29.9	36.2	104.0
SD	8.6	5.9	8.6	21.2
最高	63.0	47.0	62.0	168.0
最低	13.0	11.0	16.0	41.0
中央値	37.0	30.0	35.0	103.0

表4 学年毎の各テストにおける受験者数と平均点

	A 問題		B 問題		C 問題		全体	
	受験数	平均点	受験数	平均点	受験数	平均点	受験数	平均点
4 年生	6	31.7	6	24.7	6	30.5	6	86.8
5 年生	945	36	946	28.6	947	34.2	945	98.9
6 年生	277	44.1	277	34.5	277	43.2	277	121.8
不明	17	31.2	9	25.9	1	36	1	96.0
総計	1,245	37.7	1,238	29.9	1,231	36.2	1,229	104.0

本年度は受験者に対するフィードバックとして、各個人の得点、会場ごとの平均点、各設問の出題意図やキーワードの一覧を配布した。また、先行研究では、各個人の成績は、A 問題、B 問題、C 問題ごとの点数しか返却しなかったが、本年度は、設問ごとの正解・不正解の情報も提示した。問題を非公開とするため、具体的な問題文や選択肢、詳細な解説等を配布することはできなかったが、キーワードと合わせてその問題の正誤判定をフィードバックすることで、受験者が振り返りしやすいような状況とした。対象疾患、出題意図・キーワードの一覧と、フィードバックシートのサンプルを図2、図3に示す。

厚生労働科学研究「ITを活用した医師国家試験の採点方法の調査と検証のための研究」
2024年度 医師国家試験CBTトライアル試験

試験問題 対象疾患、出題意図・キーワード一覧

問題	対象とする疾患名	出題の意図、もしくはキーワード
A-1	骨髄片造粒症候群	骨髄片造粒症候群の診断
A-2	気管支喘息	肺の聴診所見を問う
A-3	調整性内斜視	調整性内斜視の治療
A-4	悪性黒色腫	悪性黒色腫の診断
A-5	CKD-MBD	低ナトリウム血症の鑑別
A-6	アトフィラキシー	ラテックスアレルギー症候群
A-7	Meckel憩室	Meckel憩室の組織診断
A-8	レジオナリ肺炎	レジオナリ肺炎でみられる電解質異常
A-9	急性心筋梗塞	急性心筋梗塞後合併症を問う
A-10	尿管結石	尿管結石でみられる症状
A-11	肥厚性心臓病	肥厚性心臓病をきたしやすい原因病
A-12	喉頭乳頭腫	喉頭乳頭腫の診断
A-13	血管炎候群、好酸球性多発血管炎性肉芽腫症	触知可能な紫斑(palpable purpura)をきたす疾患
A-14	二次性高血圧	二次性高血圧の原因、聴診時呼吸音異常
A-15	群発頭痛	群発頭痛
A-16	ビタミンB12欠乏症	ビタミンB12欠乏症の診断
A-17	転移性骨腫瘍	転移性骨腫瘍の診断
A-18	黄斑円孔	黄斑円孔の診断
A-19	先天性肺気腫	先天性肺気腫の症状を問う
A-20	嚥下嚥下石	嚥下嚥下石の診断と嚥下嚥下石の合併症

図2 対象疾患、出題意図・キーワード一覧

医師国家試験 CBT トライアル試験

実施日： 2024年 mm 月 dd 日
会場： ○○○○大学
受験者数： xx 名

受験番号： 24391001

(学内データ)

A 問題平均点	B 問題平均点	C 問題平均点	全体平均点
<> / 75	<> / 50	<> / 75	<> / 200

(個人データ)

A 問題		B 問題		C 問題		合計	
点数	順位	点数	順位	点数	順位	点数	順位
23	13	20	13	24	10	67	13

(個別成績)

A 問題				B 問題				C 問題							
A-1	-	A-26	-	A-51	○	B-1	○	B-26	-	C-1	+	C-26	+	C-51	○
A-2	○	A-27	○	A-52	-	B-2	-	B-27	○	C-2	+	C-27	+	C-52	-
A-3	○	A-28	○	A-53	○	B-3	+	B-28	○	C-3	+	C-28	+	C-53	+
A-4	○	A-29	-	A-54	○	B-4	○	B-29	-	C-4	+	C-29	+	C-54	-
A-5	-	A-30	-	A-55	-	B-5	-	B-30	-	C-5	+	C-30	○	C-55	○
A-6	-	A-31	-	A-56	○	B-6	-	B-31	-	C-6	+	C-31	-	C-56	○
A-7	-	A-32	○	A-57	○	B-7	+	B-32	-	C-7	○	C-32	-	C-57	○
A-8	-	A-33	-	A-58	-	B-8	-	B-33	○	C-8	-	C-33	○	C-58	-

(○)：正答 (－)：誤答

図3 フィードバックシートのサンプル

2) 大学での実施

(1) 申込管理、フィードバック

大学で実施する場合は、TAO を使って外部サーバに接続して受験することを前提にすることは

先行研究からも変更はなく、前述のようにシステムの操作画面における変更が大きな変更点となった。そのため、これまでの CBT トライアル試験に参加実績のある大学には基本的にメール等の事前通知および連絡での対応とし、大学のスタッフのみで会場準備、当日の運営を行った。一方、本年度からはじめて参加する大学、あるいは人数が大幅に変化した大学等、実施に不安を抱える大学に対しては、事前の連絡対応合わせ、可能な範囲で現地での運営支援も行った。

また、これまでの先行研究では、Google フォームを活用した参加申請の後、メールにて複数回のやり取りを通じて受験者数の確定やアカウントの案内等を行っていたが、より作業を簡素化するための方法として、申請があった段階で該当大学へ案内する事項や試験実施の直前に案内する事項など、参加申請から実施までの手順を整理するとともに、大学へ案内する実施マニュアルを作成し、効率的に実施できるような仕組みも検討した。

先行研究では TAO の契約の関係上、同一日時に受験可能な最大人数が限られており、受験者や受験日の調整が必要となる状況も生じていたが、本年度は使用上限がなくなったため、基本的には希望日程でそのまま受験可能な状態として受け入れることが可能となった。

(2) 当日の運営

各大学での CBT 試験の運用に際して、トライアル期間を通して、どの大学も大きなトラブル等はなく実施することができた。TAO の設定として、各試験問題の実施時間は固定としたが、開始時間や休憩時間は実施大学により柔軟に対応した。このため、例えば午前中に A 問題・B 問題までを済ませて昼休みとするケース、午前中は A 問題のみを実施して昼休み後に B 問題・C 問題を行うケースなどであった。

なお、個別のシステム動作に関する操作性についてはアンケートの回答としても得られている意見があるため、後述する。

3) 会場形式での実施

(1) 申込の仕組み

会場形式を実施する場合は、各大学に学生への

周知のみを依頼した。受験を希望する学生は、各個人で指定した申込フォームより参加登録を行い、研究事務局が申込受付、および問い合わせ対応等の受験者管理を行った。

また、受験終了後は、事務局から受験者個人に対し、個別のフィードバックシートを受験者が登録したメールアドレス宛に返却した。さらに、会場形式での実施にあたり学生に周知の協力をいただいた大学の担当教員に対し、事務局から当該大学を含めた、会場全体としての受験者数や平均点等の集計情報を返送した。

(2)当日の運営

本年度初めて会場形式で実施すること、さらには外部会場を借りた場合の費用負担を考えて、研究事務局のある日本医療機能評価機構の多目的会議室を使用し、ノート PC をレンタルし試験会場として準備し実施した。また、ノート PC のスペックとして、動画や画像を含む問題であるために見やすいサイズのものを選定した。CBT 試験で使用する PC のスペックや画面サイズ等については、法務省が公開している司法試験の CBT 化に関する資料なども参考とした。

当日の受験者は、20 名程の人数であり、会場内の Wi-Fi 環境を使用し、ネットワークにつないで TAO へアクセスして実施したが、大きなトラブルもなく実施することができた。また、TAO の機能であるプロクター機能を使って、試験開始、終了の操作を試験的に使用したが、試験システムの動作にも大きな影響を与えることなく問題なく実施した。



図4 会場形式の様子(会場:日本医療機能評価機構)

4)費用の試算

(1)CBT システム使用料

本研究の3年間での医師国家試験 CBT トライアル試験を想定し、TAO システムの使用料金を試算した。本研究では、あくまでも試験的な実施であるため、詳細の使用をカスタマイズし構築するのではなく、TAO サービスのパッケージとして使用した。そのため、試験実施に必要なサーバ費用等も含んでいる。

TAO の設置に伴い、初年度は初期開設費用として、25 万円程度増えるが、概ね年間で700万円～800 万円程度掛かる。なお、契約時の為替相場により、価格は変動する(表5)。

表5 CBT システムの費用概算

初期費用(1年目)	金額 (円)
TAOクラウド:CBTトライアル実施環境兼 オーサリングサーバ(年間)	7,000,000
初期開設費用	250,000
その他	125,000
合計	7,375,000

2年目以降(年間)	金額 (円)
TAOクラウド:CBTトライアル実施環境兼 オーサリングサーバ(年間)	7,000,000
合計	7,000,000

(2)外部会場使用料

CBT トライアル試験の外部会場での実施のため、都内の多目的会場で、100 人規模(1 部屋)の試験を実施する場合の試算を行った。

その結果、合計で約 1,160 万円の経費が必要となるが、その約 9 割はインターネット環境の構築費用である。動画・音声を使用した CBT 試験を想定しているため、ある程度の動画等同時アクセスに耐える環境にするためには、約 900 万円の費用が必要となる(表 6)。

今年度のトライアル試験では予算の関係上、インターネット環境を新たに構築して実施する余裕がないため、既にある程度のインターネット環境が整っている、医療機能機構の多目的会議室を使用した。

表6 会場費用概算

項目	単価	数量	単位	金額
会場借料	試験会場	180,000円	1 日	180,000円
	受験者控室	400,050円	1 日	400,050円
	スタッフ控室	150,660円	1 日	150,660円
ノートPC	7,500円	100 台		750,000円
ヘッドセット	1,500円	100 台		150,000円
電源タップ ※各机2口分(机1台あたりの価格)	3,500円	50 式		175,000円
その他 (レイアウト変更料、配送回収料等)	300,000円	1 式		300,000円
インターネット環境構築 (回線帯域1G/帯域確保100Mbps) ※短期利用でも1年以上の利用が必須	9,500,000円	1 式		9,500,000円
合計				11,605,710円

※会場は1日(8:00から17:00まで)の使用想定で算出

2. CBTトライアル試験の評価

1) アンケート結果

(1) 学生

受験者に対して実施したアンケートの結果を設問別に示す。

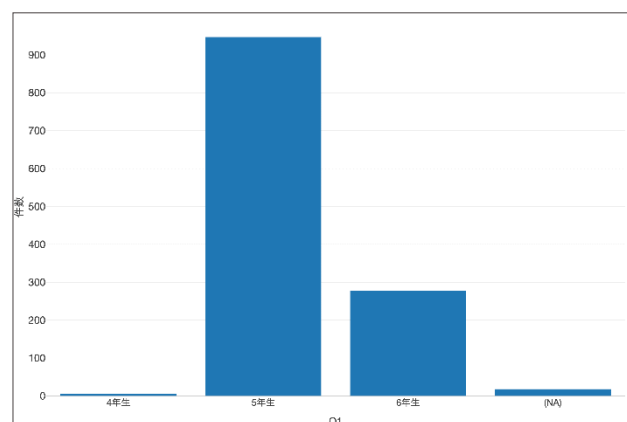


図5 学年

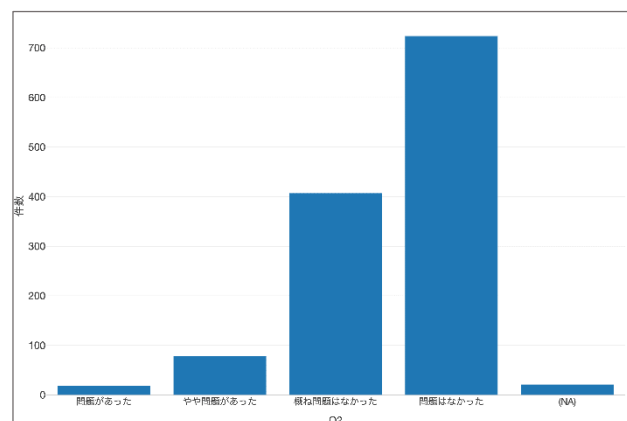


図6 操作性

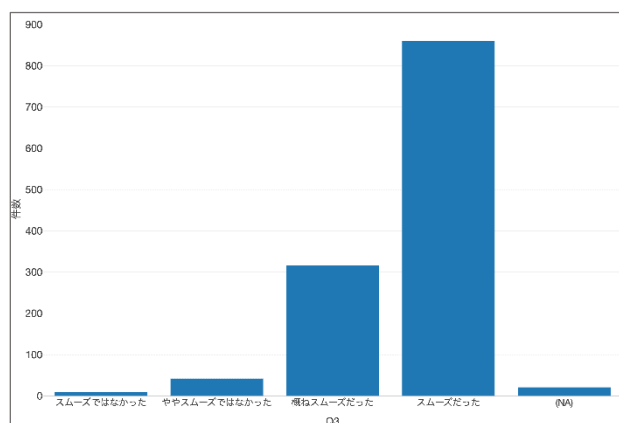


図7 動画ファイルの動作

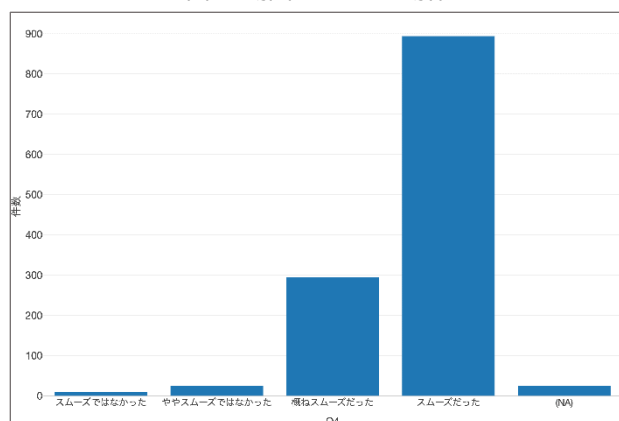


図8 音声ファイルの動作

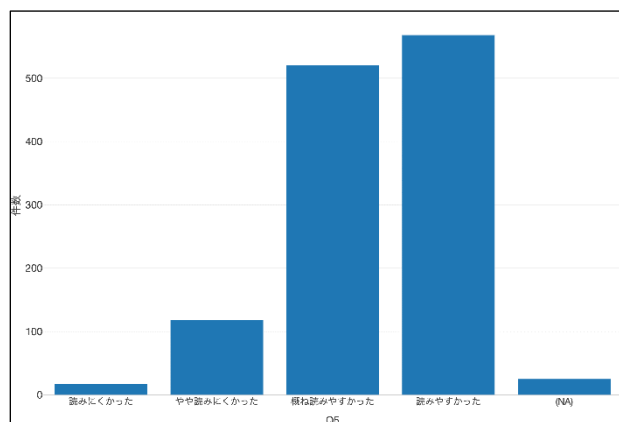


図9 文章の読みやすさ

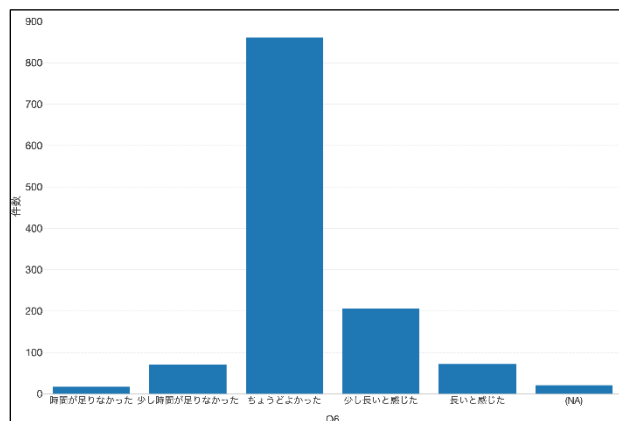


図10 試験時間

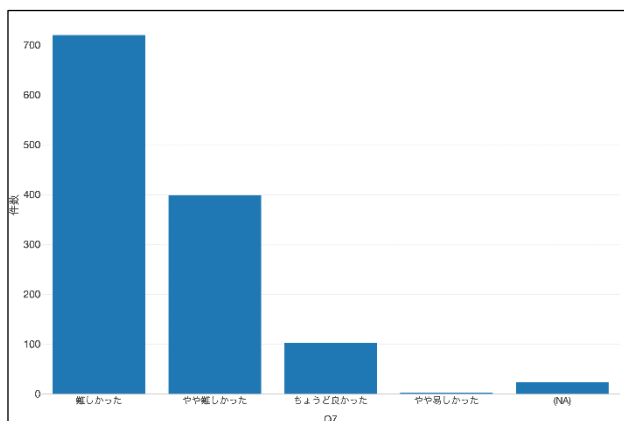


図11 難易度

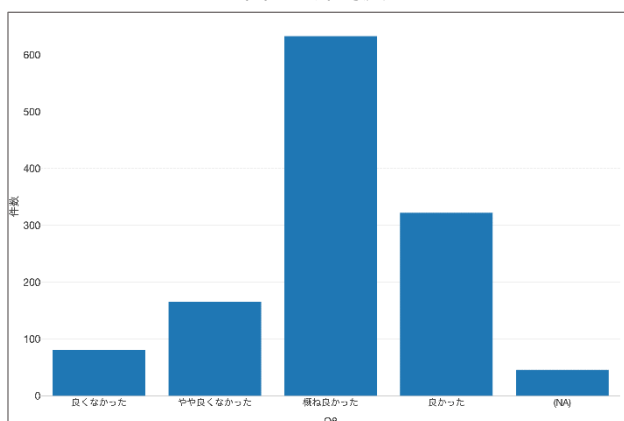


図12 CBT 試験を受験して

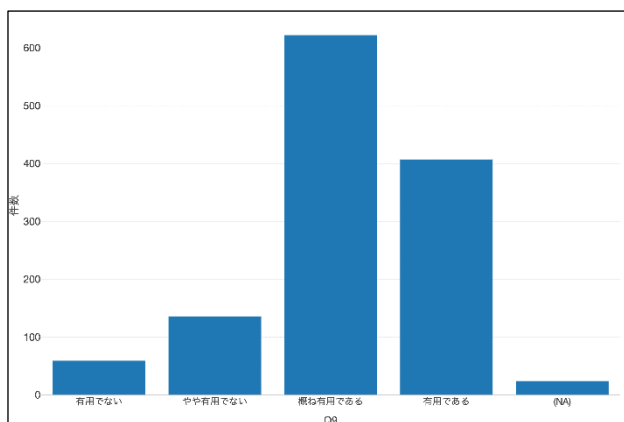


図13 マルチメディアファイルの有用性

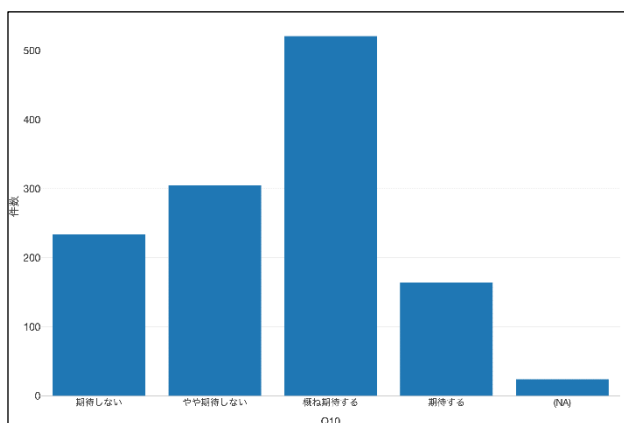


図14 医師国家試験 CBT 化への期待

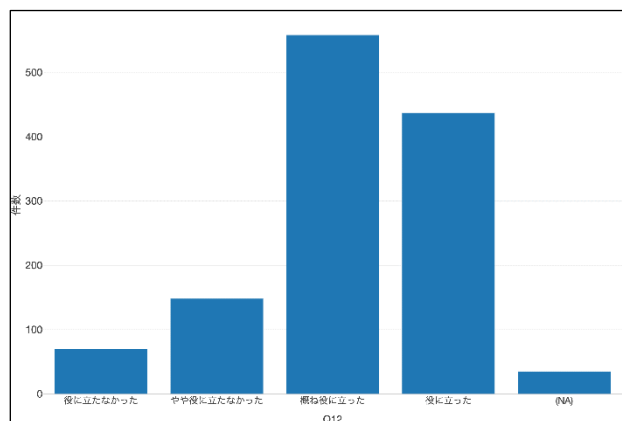


図15 メモ用紙・筆記用具

回答者の属性は受験者全体として 5 年生が 8 割弱と最も多かった。一方、1%未満ではあるものの 4 年生の受験者が含まれていた(図 5)。

CBT の全体的な操作、動画や音声ファイルの再生に関するスムーズさは、半数以上が問題なしと答えており、概ね問題ないとする回答を含めると 9 割以上が問題なく受験できていたと考えられる(図 6、図7、図8)。

一方、文章の読みやすさについては、読みやすいと回答した受験者が半数未満であり、概ね読みやすいとした回答を含めると 9 割に満たなかった(図 9)。

試験時間については、7 割が「ちょうどよい」と回答しており、「少し長い」「少し足りない」を含めたものが全体の 9 割であった(図10)。

先行研究でも類似の傾向はみられたが、試験の難易度は、学年を問わず「難しかった」が最も多く、「難しかった」「やや難しかった」は 9 割を超えた(図 11)。

CBT を受験して良かったか否か、マルチメディア問題が有用であるか否か、という問には、いずれも「概ね良い」が最も多く、「良い」「概ね良い」は 8 割程度であった(図 12、図13)。

一方、今後の医師国家試験の CBT 化に対する期待に関する問では、「概ね期待する」が最も多いものの、「やや期待しない」「期待しない」が 5 割弱であった(図14)。

CBT トライアルで有用であった機能と、配布したメモ用紙や筆記用具の有効度合いからの結果から、特に、誤答肢を削除する機能や確認用のフラグ機能が有意義であったこと、メモ用紙というア

ナログツールを有用と考える受験者も多数いたことが示された(図15)。

一方で、CBT への期待で「やや期待しない」「期待しない」と回答した受験者について、他の設問における自由記述を確認したところ、以下のようなコメントがみられた。

<CBT の実施形態(試験の運営)に関する意見>

- ・ 共通の問題でもパソコンを使うことで不平等感が生まれる。個々で出題される問題や順番が違っていると不平等である。
- ・ 公平性や平等な試験機会の確保が求められるため、パソコン上で実施は不適切。
- ・ 国家試験は二日間に渡るため、パソコンの画面を見続けて試験をするのは身体的にも苦痛が大きいと感じた。

<CBT のシステムに関する意見>

- ・ 画面の問題文や選択肢に直接メモの書き込みができないので紙の方が解きやすい。
- ・ 画面の明るさ設定、音量設定、画像の拡大ができると良い。画像の解像度を上げてほしい。
- ・ 動画を停止しながら見ることができる機能を付帯してほしい。
- ・ 操作方法やボタンの説明について事前に慣れてから受験したかった。

<CBT の内容・レベルに関する意見>

- ・ 問題が非常に難しかった。回答がもらえないとやる意味が全くない。
- ・ 動画・音声を使った問題は良いが、それに見合った実習、講義が行われていたとは言えない。
- ・ 臨症的な問題は、実習中に学んだ経験を活かす意図があると思うが、学生によって経験する臨床現場が異なるので公平性がない。

(2)教職員

CBT トライアル試験の参加大学の担当者に試験実施後のアンケートを依頼したところ、36 大学 49 名より回答が得られた。回答者属性の内訳は、医学教育担当教員が 26 名、講師が 1 名、試験担当教員および事務が 1 名、大学事務部門担当者が 20 名、大学病院職員が 1 名であった。各設問の回答結果について、図 16～図 19 に示す。

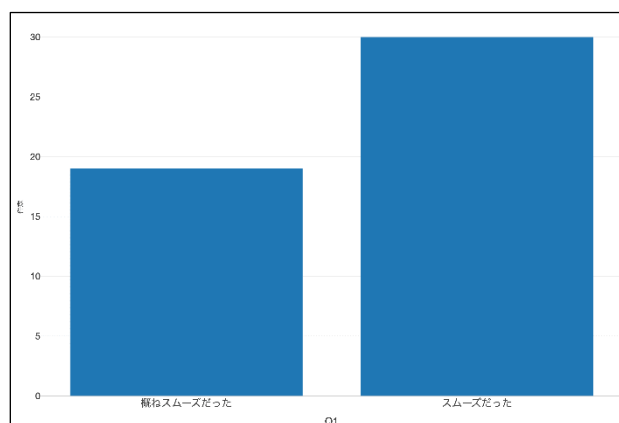


図16 CBTトライアル試験の準備や運営

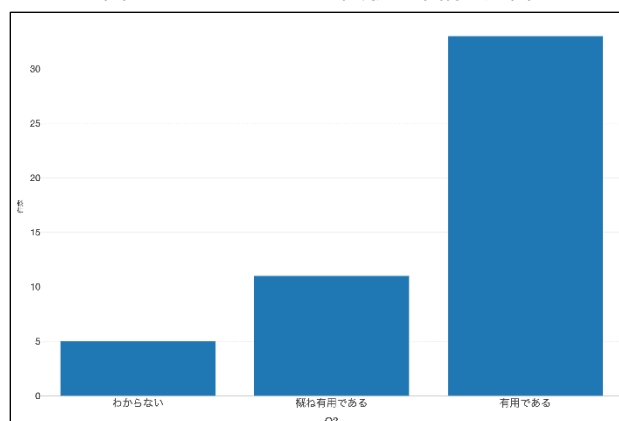


図17 マルチメディアファイルの有用性

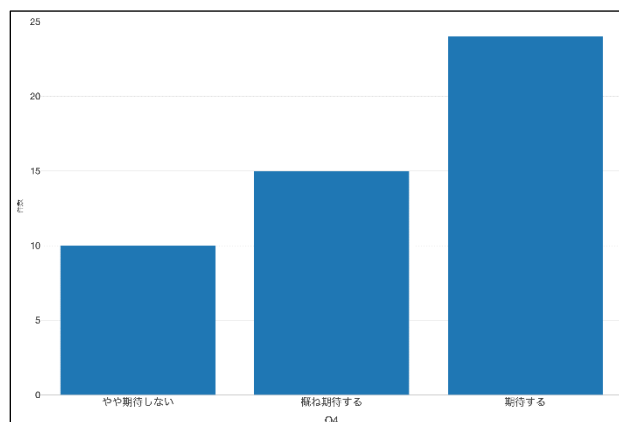


図18 医師国家試験 CBT 化への期待

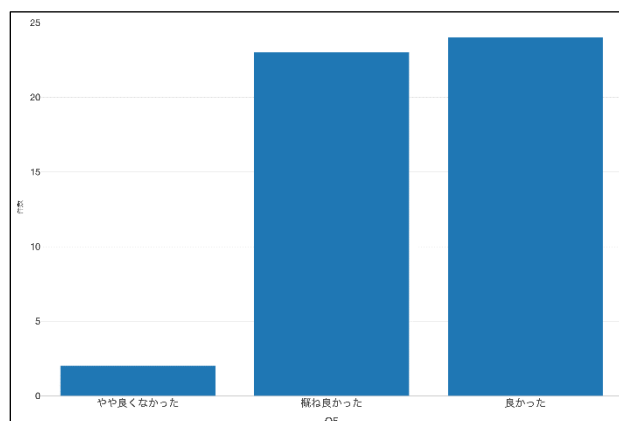


図19 CBT 試験を受験して

CBTトライアル試験のスムーズさについて、6割程度が「スムーズであった」と回答しており、否定的な意見はなかった(図16)。

動画・音声付き問題の有用さについては、6割程度が「有用である」と回答していた。「わからない」との回答が5件あったが、否定的な意見はなかった(図17)。

医師国家試験のCBT化については約半数が「期待する」と回答したが、「やや期待しない」の回答も2割程度あった(図18)。

CBTトライアルを実施して良かったかについては、約半数が「良かった」、また約半数が「概ね良かった」と回答したが、「やや良くなかった」という回答も2件あった(図19)。

CBTに対して「やや期待しない」と回答した回答者からは、以下のような自由記述のコメントがあった。これらの大半は「CBTの実施形態(試験の運営)に関する意見」として、貴重な意見として考えられる。

- ・ 共用試験 CBT のように各大学での実施となると、厳正な状況下での試験運営は大学側の教員および職員の負担と責任が非常に大きい。
- ・ 各大学で実施することは、教員・事務職員ともに人員削減が進められる中で、さらなる業務負担が増すことに懸念がある。
- ・ 各大学で実施すると、大学として PC を設置した情報端末室を整備・維持できるかが不透明。
- ・ BYOD 環境下では国家試験レベルの CBT を厳密に運用することは不可能。
- ・ 実施場所や実施方法によっては、運用面で受験者が公平に試験に臨めるか心配。
- ・ CBT 化によって試験問題が公開されなくなると、受験生の負担は大きくなる。

2) 試験問題の評価(IRT)

(1) A 問題(N=1245)の分析結果

全75項目の通過率の平均は0.503、IT 関連の平均は0.252であった。通過率が低かった項目は A35(0.050)、A19(0.108)、A21(0.149)、高かった項目は A65(0.933)、A2(0.899)、A53(0.880)であった。IT 関連が大きかった項目は順に A39(0.548)、A14(0.500)、A16

(0.467)、A59(0.445)であり、0.2を下回った項目は A53(-0.036)、A74(0.011)、A13(0.048)など26項目あった。信頼性の指標であるクロンバックの α 係数は0.86であった。

2PLMによるA問題の項目母数の推定値を図1に示した。プロットされている文字は項目番号を表す。困難度母数は概ね ± 3.0 の間にまんべんなく散らばっている。通過率が0.2に満たなかったA21、A20、A06の困難度母数の推定値は大きく、通過率が0.8を超えていたA66、A63の困難度母数の推定値は小さく、古典的テスト理論に基づく結果とIRTモデルによる推定結果との整合性が見られた。一方で、A41とA53に注目すると、通過率はそれぞれ0.698と0.880で正答率に差があるように見えるが、2PLMで推定された困難度母数の値は-2.636と-2.403であった。これは、A41、A53はいずれも $\theta = -2.5$ 前後の能力値の受験者であれば5割の確率で正当できる項目であることを意味する。IRTの最大の利点は、項目の特徴を表す困難度および識別力と、受験者の能力とを分離して評価できるという点であり、その結果として、A41とA53は項目自体の難しさに大きな差異はないことが示された。また、A14、A39、A59は中程度の困難度と高い識別力を備えた良問と言える。このうちA39とA59は画像を使用した問題であった(図20)。

図20 2024年度A問題の項目母数(49項目)

(2)B 問題(N=1238)の分析結果

全 75 項目の通過率の平均は 0.600、IT 相関の平均は 0.265 であった。通過率が 0.2 未満だったのは B32(0.064)、B31(0.183)、B19(0.191)の 3 項目、0.9 以上であったのは B01(0.951)、B28(0.930)、B37(0.929)、B25(0.904)の 4 項目であった。IT 相関の最大値は B44 の 0.402 であり、0.2 を下回った項目は B39(0.046)、B38(0.127)、B37(0.148)など 13 項目あった。クロンバックの α 係数は 0.72 であった。

2PLM による B 問題の項目母数の推定値を図 2 に示した。2024 年度に実施した B 問題では、困難度母数の小さな、すなわち簡単な項目の方識別力が高い傾向が見られた。ただし、一般に識別力母数の値は概ね 0.3~2.0 の間で推定されることを考慮すると、識別力母数の推定値が 1.0 以上だったのは B25(1.279)、B49(1.232)、B42(1.047)の 3 項目のみであり、概して識別力が低い項目が多かった。なお、B25 は動画および音声を使用した項目、B42 は連問の 2 問目、B49 は連問の 1 問目であった。また、B32 は通過率 0.064 と非常に正答率が低い項目であり、困難度母数の推定値も最大となったが、識別力母数の推定値が比較的高いことから、難しい項目であっても一定程度の識別力を持つ項目が存在することが確認された(図21)。

(3)C 問題(N=1231)の分析結果

全 75 項目の通過率の平均は 0.483、IT 相関の平均は 0.256 であった。通過率が低かった項目は C22(0.045)、C32(0.070)、C53(0.074)、C50(0.084)、C29(0.089)、高かった項目は C65(0.929)、C45(0.883)、C10(0.877)であった。IT 相関が大きかった項目は順に C48(0.542)、C31(0.425)、C72(0.407)であり、0.2 を下回った項目は C33(0.023)、C29(0.028)、C50(0.036)など 25 項目あった。クロンバックの α 係数は 0.81 であった。

2PLM による C 問題の項目母数の推定値を図 3 に示した。C 問題では、困難度母数の推定値が極端に小さな項目の識別力は高くなりにくいという一般的な傾向が見られた。困難度母数の推定値が最も大きかった項目は C59(2.691)、最も小さかった項目は C45(-2.753)であった。識別力母数の推定値が大きかった C48(1.786)、C72(1.487)、C19(1.232)の困難度母数の値はそれぞれ-0.194、-1.226、-1.057 であり、これらは能力値の低い受験者を取りわけよく識別できる項目であることがわかる。このうち、C48 と C72 は画像を使用した問題で、C72 は 3 連問の 1 問目であった(図22)。

図22 2024 年度 C 問題の項目母数(50 項目)

図21 2024 年度 B 問題の項目母数(37 項目)

3. CBT 問題の作成

1) CBT 作問

今後の CBT 問題のプール化を見据え、各専門領域の作問者の協力を得て、新たな CBT 問題を医師国家試験のブループリントに準拠して作問し、2024 年度は、必修の基本的事項 65 問(一般問題 25 問、臨床問題 40 問)、臨床総論 157 問(一般 53 問、臨床 104 問)、臨床各論 141 問(一般 15 問、臨床 126 問)であり、総作問数は 363 問となった。そのうち動画・音声の素材を使用した問題は、必修の基本的事項で 18 問、臨床総論で 34 問、臨床各論で 21 問であり、計 73 問であった(表 7)。

これらの中から、ブラッシュ 2024 年度の CBT トライアル試験に使用する 200 問の CBT 問題を 1 セット作成するため、後述の研究協力者(ブラッシュアップメンバー)4 名と研究分担者(岡崎、伴、浅田、久保、松山)により、総合的に検討し 200 問を選定した。

表7 2024 年度作成した CBT 問題

		動画・音声		合計
		あり	なし	
必修の 基本的事項	一般	6	19	25
	臨床	12	28	40
臨床総論	一般	4	49	53
	臨床	30	74	104
臨床各論	一般	3	12	15
	臨床	18	108	126
合計		73	290	363

2) 作問者およびブラッシュアップメンバーの構成

CBT 問題の作問は、今後、数多くの問題数が必要となること、さらには作成した問題の定期的な見直しが必要となることから、計 33 名の各専門領域の作問者、およびブラッシュアップメンバーによる CBT 問題作問体制を構築した(表 2)。また、作問した問題を客観的に評価するため、CBT 問題のブラッシュアップ作業を、研究分担者(岡崎、伴、浅田、久保、松山)と 4 名の研究協力者(ブラッシュアップメンバー)により実施し、本年度に作成した CBT 問題および先行研究蓄積された問題等の評価・分析、

今後の CBT 問題のプール化について議論した。

表8 作問者およびブラッシュアップメンバー一覧

	氏名	主な領域
<作問者>		
1	阿江 竜介	公衆衛生学
2	石原 あやか	総合診療
3	磯部 真倫	産婦人科
4	猪田 宏美	薬学
5	伊藤 真人	耳鼻咽喉科
6	及川 沙耶佳	救急医学
7	小川 尊資	皮膚科
8	兼田 裕司	一般・消化器外科
9	菅野 武	消化器内科
10	喜多村 洋幸	歯科口腔外科
11	黄 世捷	循環器内科
12	才津 旭弘	整形外科
13	笹原 鉄平	感染症
14	鋪野 紀好	総合診療
15	鈴木 寛正	産婦人科
16	鈴木 真紀	総合診療
17	田鎖 愛理	精神科・産業医学
18	田島 敏広	小児科
19	田中 淳一	呼吸器内科
20	林 幹雄	総合診療
21	福岡 敏雄	救急医学
22	松本 健吾	形成外科
23	松本 卓子	呼吸器外科
24	三浦 聖子	一般・消化器外科
25	三原 弘	消化器内科
26	野村 理	小児科・救急医学
27	矢吹 拓	総合診療
28	山本 祐	総合診療
29	早稻田 勝久	循環器内科
<ブラッシュアップメンバー>		
1	見坂 恒明	総合診療
2	鈴木 昌	救急医学
3	西川 正憲	呼吸器内科
4	宮崎 泰司	血液内科

3) CBT 問題の作成手順の標準化

CBT 問題の作問は、先行研究では、作問者が PowerPoint ベースで作成し、その問題を TAO に入力するという方策をとっていた。この場合、PowerPoint という普段から使い慣れているツールが利用できる利点はあるものの、入力の手間が二重に生じること、ブラッシュアップ等を行う際にバージョン管理が困難になりがちなことなどの課題

がある。このため、作問の段階から標準化されたフォーマットを使用しオンライン上で入力することができ作業効率を高められる方法の検討を開始した。

出題までの作業効率のみを考えると、TAOに直接入力することが最短と考えられる。一方、作成した問題のブラッシュアップ、プール問題としての保存、利活用情報のデータベース整理などを考えると、TAO 以外のシステムで作成した後、QTI (Question and Test Interoperability)形式でインポートする方が汎用性を高くすることができる。これは将来的に、TAO というシステム以外の利用を検討せざるを得ない場合になったときも、移行を容易にするために意義のある方策となる。

本年度はその第一段階として、Moodle のデータベース機能を用いた作問システムを整備した。Moodle はオープンソースで提供されている LMS (Learning Management System)であり、主に非同期型での学習等に利用されることが多い。Moodle には様々な機能が存在するが、今回はその中の「データベース」機能を活用した作問システムを用意した。Moodle のデータベース機能は、文字通り Moodle 内部に簡易的なデータベースを設定するものである。入力された内容は CSV ファイルとしてダウンロードが可能である。また、Moodle 4.3 より、画像や動画等のファイルを添付した場合も、ファイル名と紐づけてダウンロードが可能となった。

作成した入力フォームでは、左右に入力フォームが存在していることが分かる。左側を初回入力、右側を改訂用の入力欄として、使い分けられるようにした。また、誤操作を防ぐため、入力に際してはチェックボックスの ON/OFF を利用するようにした。この他、正解選択肢の入力に加え、解説や備考を記載できる欄を追加した(図23)。

動画素材等を活用した CBT 問題の作成方法について、マニュアルを作成するとともに、統一した作問の入力フォームを提供することで、一定程度標準化した CBT 問題の作成を可能とすることができた。なお、完成した CBT 問題作成マニュアルは、参考資料として別途掲載する。

図 23 作成した入力フォームの一部

4. 国内国家資格試験における CBT 試験導入の状況

1) 法務省

法務省では、「デジタル社会の実現に向けた重点計画」(令和 5 年 6 月 9 日閣議決定)において、「司法試験および司法試験予備試験のデジタル化」「CBT 方式による試験について、令和 8 年に実施する試験からの導入に向け、システムの設計・構築等を進める」と言及されたこと、および「経済財政運営と改革の基本方針 2023」(令和 5 年 6 月 16 日閣議決定)において、「司法試験のデジタル化の推進」を図ることが示されたことから、令和 8 年度より、司法試験および司法試験予備試験において CBT 形式での実施が予定されている。

CBT システムを独自で開発し、司法試験の特徴である長文での記述回答のデジタル化などを実現している。

2) 国土交通省

国土交通省が実施する、航空従事者技能証明等学科試験は、年間の受験者数は、5,000 名から 6,000 名であり、コロナ過での資格試験の受験状況から専門学校や航空会社からの要望により、2023 年 11 月より CBT 化された。また、これまで試験問題を公開していたが、CBT 化に合わせて非公開となった。試験問題には動画・音声等は含まれておらず、運営については CBT 業者に委託(毎年競争入札)して、CBT 業者が運営する全国の会場で実施している。

CBT 化する前に、紙で試験を行っていたときは、

年2回から6回(資格の種類等によって異なる)実施されていたが、CBT化されてからは、年6回の試験期間(1期間10日間程度)設け、受験者が受験日や、会場を希望して選択でできる。

CBT問題は約7,000題をプールしており、IRTを取り入れ、受験者間のレベルの差や、受験者が同じ問題を受けることがないように対応をしている。また、問題を非公開として運用しているが、現状、問題の漏洩など大きな問題は起きていない。その理由として、航空従事者技能証明等学科試験は、医師国家試験と異なり、予備校がないこと、またこれまでの過去問が多く公開されていることがあげられる。

3) 総務省

総務省が実施する、無線従事者国家試験では、試験日程や試験場所の拡大による受験者の利便性の向上を目的として、令和4年からCBT方式の国家試験を開始した。

試験問題の過去問はCBTのため公開されておらず、例題およびその回答が公開されている。

CBT方式の試験では試験日の3日前までの手続きにより試験日程・会場の変更が可能である。ただし、最初の申請の日から120日以降の試験日は選択できない。また、試験日の3日前まではキャンセルも可能である。

試験結果は概ね2週間以内に電子メールで連絡されることとされている。

4) 経済産業省

経済産業省が実施する、情報セキュリティマネジメント試験および基本情報技術者試験では、これまで試験会場で配布されるマークシートに記入する形式で実施していたが、令和2年4月の試験が新型コロナウイルス感染症の影響により中止となり、さらに以降の試験会場の確保が困難となったことから、令和2年12月以降からCBT方式が正式に導入された。試験の実施方法がマークシートからCBT形式に変更されたのみで、出題形式および範囲は従来の形を踏襲し、試験の日程は上期(6月)と下期(12月)の年2回実施されている。

また、試験問題は持ち帰りできず、非公開とし、受験者には、受験時に試験問題の第三者への開

示(漏洩)や開示(漏洩)した場合の措置について周知徹底するなどの漏洩対策を行っている。

5) 文部科学省

文部科学省では、国家資格試験ではないが、児童生徒が学校や家庭において学習やアセスメントができるCBTシステム(MEXCBT)を開発し、令和3年12月から、運用を開始し、児童生徒1人1台端末を活用した学習を取り入れており、CBTシステムを活用して、全国学力・学習状況調査を実施している。

D. 考察

1. CBT試験の実施運営とシステム

CBTトライアル試験の実施は、インターネットを介したCBTシステムTAOを活用し、画像や音声、動画等のマルチメディアを利用したより実臨床に近い試験問題を作成して、全国52大学(1,248名の受験者)で概ね円滑に実施することができた。また、大学での実施のみならず、会場形式として大学以外の会場で1回実施するとともに、トライアル試験の運営に際し、プロクター機能や公開セッション、Moodleでの作問管理など、試験運用を円滑に進めるための技術的な新しい取り組みも取り入れたことで、これまでの大学での実施における課題や改善策に加え、さらに幅広い運用にあたっての課題や改善策などの知見を得ることができた。

システムの利用面について、CBTの全体的な操作、動画や音声ファイルの再生に関するスムーズの設問と比較し、文章の読みやすさについてはやや低い結果となっていた。この理由として、動画や音声といった単体のファイルを再生することと異なり、「問題の文章の読みやすさ」という問いから、受験者は「文字の読みやすさ(行間やフォントサイズ)」「画像の見やすさ」「全体的な画面配置」などをすべて網羅的に受け止めて判断した可能性がある。今回、問題の表示形式を統一化することを検討したが、臨床問題のように問題文が長いもの、画像や動画が複数含まれるものなどについては、スクロールしないとすべてが表示されないという状態になっていた。

今回、CBTトライアルにおいては、システムとし

での動作検証を行う以外に、医師国家試験の内容や難易度を検討することを念頭に、通常の医師国家試験において正答率が高くなり得る問題を意図的に外しているという作問を行っている。全体として難易度に対しては難しいという意見が多かった理由の1つとして考えることができる。

今後の医師国家試験のCBT化に対する期待については、ほかと比べて回答結果が低い値となっていた。この理由について、個別の自由記述を付与しなかったため、厳密な理由については追うことが困難な状況である。しかし、「やや期待しない」「期待しない」と回答した受験生であっても「CBTを受けてよかった」と回答しているケースや自由記述で肯定的な内容を書いているケースもあったことから、一概に期待する度合いが低いとも考えづらい。1つの可能性としては、この「期待する」という問い方に対して、「実現する可能性として、期待するか否か」「自分たちが受験するにあたって、CBT化してほしいか否か」など、複数の受け止め方が起こり得ることである。この場合、回答者によって異なるイメージを持っていることとなり、回答の精度に影響が出ている可能性が否めない。

ツールの利用に関するアンケート結果からは、TAOの機能を十分に利用して望んでいる受験者が多いことも分かる一方、メモ用紙というアナログツールの必要性も示唆されていることがうかがえた。紙というツール以外に、タブレット端末等を利用することも将来的には一案として考えられる可能性がある。

2.CBT問題の作成と体制

先行研究で培った方法を活用しCBT問題の作問は概ね円滑に進行した。マニュアルが整備され、Moodle提出フォームとの連携により、作問および提出にかかる負担が軽減された。動画については、個人情報保護の観点から、編集により個人が特定されない形式とする、あるいは模擬患者を用いて異常所見を再現するなど、適切な方法でマルチメディア素材が活用された。

ブラッシュアップ作業においては、複数のメンバーによる確認を通じて、疑義が生じにくく統一感のある記述へと整備された。また、先行研究で蓄積された臨床問題の正答率や識別力等のデータを活

用し、受験者の能力を適切に識別しうる問題へと改良することが可能であった。

特に、過去にCBTトライアルで出題され、正答率や識別力が低かった問題については、シナリオやマルチメディア素材を再利用しつつ、新たな設問構造へと再設計することで、新作問題としての活用が可能であることが確認された。

以上の結果から、過去の解答結果に基づき、受験者の能力を的確に評価できる問題を厳選することで、試験問題数を削減して受験時間の短縮を図りつつも、医師国家試験としての妥当性を維持することが可能であると考えられた。この観点から、問題の非公開化を前提とした項目反応理論(IRT)を用いた医師国家試験CBTの実施は、今後の有用な方向性の一つであると示唆される。

3.CBT問題のプール化、IRT

本年度実施したCBTトライアル試験の回答データに対し、これまでと同様に古典的テスト理論に基づく項目分析と、IRTの2PLMによる分析を行った。CBTトライアル試験は2021年度から毎年実施しており今回で4回目となるが、2024年度はA問題、B問題、C問題ともに最も平均点が低かった。項目ごとの通過率の平均値も小さくなり、IT相関が0.2未満隣IRTによる分析が適用できなかった項目の数はこれまでで最も多くなった。IRTによる分析結果においても、前年度までと比較して識別力母数の値が小さく推定される傾向が見られた。これらの結果が、前年度までと比較して2024年度の受験者では学年の構成比が大きく異なるなどの理由で能力が変化したことによるのか、あるいは試験を構成する問題の難易度など項目の特性が変化したことによるのか、要因を検証する必要がある。

さらに、今後の課題として、等化(equating)の必要性が挙げられる。IRTによるテスト運用では、あらかじめ共通尺度に等化された項目母数をもつ多くの項目を用意して項目プールを構築することが最も肝要である。本研究においては、共通項目計画による尺度の等化を視野に入れ、年度をまたいで共通する項目を含めてCBTトライアル試験を実施している。共通項目の項目母数を用いて等化を行い、等化の手続きを経てプール化された項目

を利用して試験を実施することで、等質な複数の試験(フォーム)を作成したり、受験者の異なる試験で得られた得点を共通尺度上で解釈したりすることが可能となる。

また、将来的に項目のプール化を効率的に行っていくために、本年度はこれまでのCBTトライアル試験で出題した問題を対象として、古典的テスト理論に基づく項目分析およびIRTモデルによる分析の結果を問題作成者にフィードバックする場を設けた。1つ1つの問題に関して、問題作成者と分析者がともに分析結果と問題の内容を参照しながら、なぜそのような分析結果となったのか考え得る要因について意見交換を行った。テスト理論の観点から識別力の高い問題と低い問題、あるいは難易度の高い問題と低い問題それぞれの具体的な特徴が明らかとなり、良問の効率的かつ継続的な作成にあたって、このような議論の機会の有効性が確認された。

4. 国内国家資格試験における CBT 試験導入の状況

国土交通省の航空従事者技能証明等学科試験、総務省の無線従事者国家試験、経済産業省の情報セキュリティマネジメント試験および基本情報技術者試験において、既にCBT形式での試験運用が行われており、さらに、試験問題を非公開として実施されていることから、医師国家試験においても実装に向けた課題は多々あるが、これらの先行的に行われている国家資格試験を参考に検討を進めることで、より実現に向けた具体的な提案が可能となる。

さらに、他省庁とCBT試験の取り組み状況等について情報交換を行いながら進めることで、省庁の垣根を越えた我が国のCBT国家資格試験の早期の実施が可能と考える。

E. 結論

CBT 医師国家試験の導入の実現に向けたCBTトライアル試験では、将来の CBT 医師国家試験を見据え、先行研究からの大学での実施に加えて、大学以外の会場形式での実施や、CBTシステムの試験機能の改善、プロクター機能を使った

試験運用、受験者数の制限をなくした複数大学の同日試験の実施などの新たな方法を取り入れて、全国 52 大学(1,248 名の受験者)で円滑に実施した。

先行研究からこれまでの4年間で、大きなトラブルなく概ね円滑に実施し、この CBT システムにより CBT 医師国家試験を実装することができることを実証した。また、医師国家試験が CBT 化された場合、多くの CBT 問題が必要となることから、33名の各領域の専門家からなる CBT 問題の作成・管理体制を構築し、約400問の新規問題作成、さらに先行研究で作成した約600問についてIRT分析を用いた評価を行うとともに、作問者および評価者として協議する体制を整えた。

CBT 問題のプール化や CBT 試験問題の作成体制と CBT 試験システム等について検討を行い、CBT 医師国家試験の実装に必要な具体的な課題や対応策などの知見を得ることができた。

先行研究からこれまでの4年間のトライアル試験における、受験生および大学の教職員に対するアンケートでは、医師国家試験の実現に向けて期待する意見がある一方、今後実現に向けて検討すべき課題が示された。医師国家試験の CBT 化を実現するためには、受験生および大学の教職員の理解が極めて重要であることから、これらについては引き続き改善策を検討するとともに、トライアル試験において改善策を試行する必要がある。

我が国の国家資格試験のCBT化については、欧米諸国と比較すると導入が遅れていると言われているが、他の省庁が所管する国家資格試験において既にCBT形式が導入されていること、さらに法務省で司法試験のCBT形式の導入が予定されている現状から、他省庁のCBT化の動きを参考にし、医師国家試験においても、CBT化を実現することが強く望まれる。

F. 健康危険情報

該当なし

G. 研究発表

1. 論文発表
なし

2. 学会発表

1) 第 56 回日本医学教育学会大会

オンデマンドシンポジウム：医師国家試験 CBT
化に向けた研究成果と今後の展望

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

記載すべき事項なし

I. 研究協力者一覧(50 音順・敬称略)

2025年3月31日時点

<CBT 問題作成・検討>

阿江 竜介	(自治医科大学 自治医科大学地域医療学センター 公衆衛生学部門)
石原 あやか	(大分県立病院 リウマチ膠原病内科)
磯部 真倫	(岐阜大学 産婦人科)
猪田 宏美	(岡山大学病院)
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)
小川 尊資	(順天堂大学 医学教育研究室・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)
喜多村 洋幸	(松本歯科大学病院 初診室(総合診断科・総合診療科))
見坂 恒明	(神戸大学 大学院医学系研究科 医学教育学分野 地域医療支援学部門)
黄 世捷	(聖マリアンナ医科大学 医学教育文化部門)
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)
笹原 鉄平	(自治医科大学附属病院 感染制御部)
鋪野 紀好	(千葉大学 大学院医学研究院 地域医療教育学)
鈴木 寛正	(自治医科大学 産科婦人科学講座)
鈴木 真紀	(益田赤十字病院 総合診療科)
鈴木 昌	(東京歯科大学 市川総合病院)
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)
田島 敏広	(自治医科大学 医学部小児科)
田中 淳一	(東北大学 大学院医学系研究科 医学教育推進センター)
西川 正憲	(藤沢市民病院)
野村 理	(岐阜大学 医学教育開発センター)

林 幹雄	(関西医科大学 教育センター)
福岡 敏雄	(倉敷中央病院)
松本 健吾	(社会医療法人敬和会大分岡病院 形成外科)
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)
三浦 聖子	(島根大学 医学部 救急医学講座)
三原 弘	(札幌医科大学 医療人育成センター)
宮崎 泰司	(長崎大学 原爆後障害医療研究施設)
矢吹 拓	(国立病院機構 栃木医療センター 内科)
山本 祐	(自治医科大学 地域医療学センター総合診療部門)
早稲田 勝久	(愛知医科大学 医学教育センター)

<CBT システム検討>

CROSS Jeffrey Scott (東京科学大学)

【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

