

医師国家試験の海外における CBT 試験の動向と CBT システムに関する研究

研究分担者 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

本分担研究では、医師国家試験の CBT 化に向けて必要な条件整備について研究するために、CBT 実施会社の各社について比較検討してこられた門川俊明先生(慶応義塾大学医学教育統轄センター教授)にオンラインでヒアリング調査を実施し、各社の施設、設備等の情報収集を行った。さらに、医療系大学間共用試験実施評価機構の試験信頼性向上専門部会において、試験問題分析、試験問題事後解析について検討してこられ、現在 OECD の分析官として活躍されている大久保智哉先生(The Directorate for Education and Skills, OECD)と、研究分担者の久保沙織と伴信太郎が、今後の医師国家試験の CBT 化に向けて、海外の国家試験の CBT 化の状況や、試験の実施方法・実施場所、試験体制などの課題について検討した。

A. 研究目的

我が国における医師国家試験の CBT 化に向けて、海外の国家試験の CBT 化の状況や、CBT 試験の実施方法・実施場所、試験の運営体制などについて、CBT 化を実現するために必要な条件や課題を整理することを目的とする。

B. 研究方法

1. CBT のベンダー比較

CBT 実施会社の各社について比較検討してこられた門川俊明先生(慶応大学教授)に Zoom ミーティングでヒアリングを実施し、各社の施設、設備等の情報収集を行った。

2. 医師国家試験 CBT 化の設計

医療系大学間共用試験実施評価機構の試験信頼性向上専門部会において、試験問題分析、試験問題事後解析について検討してこられ、現在 OECD の分析官として活躍されている大久保智哉先生と、河北班研究分担者の久保沙織と伴信太郎が今後の医師国家試験の CBT 化に向けて必要な条件や課題について検討した。

C. 研究結果

1. CBT のベンダー比較

1)門川先生のプレゼンテーション『CBT 実施会社の比較』

(1)CBT 実施4社の比較

2020 年当時の情報であることを留意する必要があるが、(株)プロメトリック、(株)ピアソン VUE は海外ベンダーの日本支社であり、(株)教育測定研究所、(株)CBTソリューションズは日本企業である。CBT 実施会社は4社以外にもあるが小規模である。

①(株)プロメトリック

最も規模が大きく、日本への参入時期も早い。外資系企業の日本支社で会場のセキュリティがしっかりしている。費用は高めである。

②(株)教育測定研究所

参入が最も新しい日本の企業で、規模はプロメトリックに次ぐ。直営 32 会場、委託 60 会場、貸会議室 290 会場を使って、全国 47 都道府県を網羅できる。会場の質が担保されている。

費用は、300 人程度の受験生の場合で、2 時間の試験時間、13 会場を確保して、500 万円程度の予算が必要となる。

③(株)CBTソリューションズ

色々なテスト会場を借りてそこで CBT を実施する。幾つかの学会の専門医試験の経験がある。

④(株)ピアソンVUE

プロメトリックに次いでCBTベンダーとして日本に参入した外資系企業の日本支社である。会場におけるセキュリティがしっかりしている。

表1 CBT 実施会社の比較

(門川教授提供資料より一部改変:2020年当時)

会社名	配信数	参入時期	特徴	専門医試験の実績
プロメトリック	250万/年	1991年	ETSが買収。TOEFLなど実績が豊富	総合診療専門医、麻酔など
教育測定研究所	60万/年	2019年	英検CBTなど。直営会場が43会場と多いのが特徴。	整形外科、消化器病など
CBTソリューションズ	50万/年	2009年	漢字検定、ソムリエ試験など	リウマチ、高血圧、アレルギー学会など
ピアソンVUE	40万/年	2000年		日本周産期・新生児医学会など

(2) CBT 実施会社選定にあたって検討すべきポイント

- 直営の試験会場が確保できるか。
- 試験会場を CBT 実施会社が確保するのか、試験実施者が自分で確保するのか。
- 試験会場の質は均一に保たれているか。
- 同時開催は何人まで可能か。
- 予算

2) 医師国家試験 CBT 化に向けての課題

- 教育測定研究所以外は 300 人規模でも同時開催は困難である。
- 会場を借りて行う場合、ベンダーが使っているシステムに CBT問題を入れるような形であるため、動画等が入っている問題を入れ込めるかどうかはベンダーに確認する必要がある。会社によってはフォント等の制限もある。
- 教育測定研究所の見積もり 500 万円には、システム側の費用、試験後の様々なデータのアウトプットなども含めている。この金額は交渉の結果ディスカウントされた金額である。

- 最初のデータの流し込みや動作確認等で数か月の準備期間が必要である。
- 同時開催で最大で、実績としては、全国で 10,000 人近い数で実施した経験がある。
- ピアソン VUE は、同時開催が困難であったため、見積もりまではいかなかった。見積もり価格は、CBTソリューションズが一番安価であったが、同時開催は不可とのことであった。
- 実施時期によって同時に開催できる人数は変わってくる。
- 教育測定研究所は、後発であるので実績を増やしたいと、積極的に参入を希望していた。
- 外資系の会社は強気であった。
- 問題セットの出題順を受験生ごとにランダムに出題するようなアレンジは可能である。
- 現在、各ベンダーが実施している CBT の試験は 1-2 時間の短時間の試験が殆どであり、2日間の試験を実施するためにはそのための準備が必要である。
- 会場の環境についてはパソコン教室のような会場を用意している場合もあり、その点では外資系の実績のあるベンダーの方が信頼をおける。
- 試験問題をプールする場合には、どこにプールするのかを検討する必要がある。
- 見積もり価格の細目としては、会場費、会場を管理する人員の人件費、試験本部との連絡体制等でかなりの費用がかかっているため、試験時間が5倍になっても費用が 500 万円×5になるわけではない。

2. 医師国家試験 CBT 化の設計

1) 大久保先生の『日本の医師国家試験制度の設計』についての意見

(1) 医師国家の CBT 化は必至

全世界で様々な試験の CBT 化が進んでいる現状(表2、表3)から、我が国における医師国家試験を CBT 化しないという議論はあり得ない(ナンセンス)と言ってよい。

(2) 一種類の試験で医師としての資格の有無を問うのは著しく妥当性に欠ける

一種類の試験のみでハイ・ステイクスな試験の可否を判定するのではなく、複数の評価の組み合わせ(パッケージ)で判定するのが世界ハイ・ステイクスな試験の趨勢である。過度に信頼性を追求するのは、方向性としては問題がある。

(3) 国家試験の実施場所・実施方法について

① 国家試験の実施場所

テスト・ベンダーに依頼して実施するのが現実的ではないかと思われる。

【理由】

- セキュリティの確保とコストとのバランスが良い。
 - テストの管理・実施・実施後のデータ処理の経験が豊富で安定している。
 - テスト実施場所を全国的に確保している。
- #### ② 試験の実施方法(長期的な展望を含めて)
- 医師国家試験を分野ごとにモジュール化(ex.3~6 分野)して、全体としての問題数は増やすが、モジュールごとに受験を認め、最終的にはすべてモジュールに合格したら資格を与える。
 - 受験時期は臨床実習を開始後はいつでも受験可能とする(4年次でも、5年次でも、6年次でも)。ただし、一つのモジュールの受験可能回数を制限する。USMLE は同一領域の受験は年 3 回までに制限している。こうすることによって現状のような半年~1 年間の臨床実習をしないで筆記試験に備えるという弊害を防げるのみならず、現在の医師国家試験のように相対評価をして結果的に競争試験になってしまう弊害も防げる。

2) 医師国家試験 CBT 化に向けての課題

(1) 現在のトライアルで見られる作問の問題点

- 古典的なテスト理論での IT 相関を検討してみると、2021 年度、2022 年度の A、B、C 問題の 1-3 割の問題が 0.2 を下回っており識別力が低く、これらの問題の質向上の余地があると思われる。今回の IRT の分析には、IT 相関が 0.2 以下の問

題は除外した。これらの問題が、どのような点で質の向上の余地があるのか検討すべきである。

- どのような試験でも 2-3 割はそのような問題が出てくる。5 割くらいの場合も少ない。
- 今回のトライアルでこのような作問の問題が明らかになったことは、試験問題の試行的出題をして教育測定学的に検討した上で問題のブラッシュアップをする重要性を示している。
- 新作問題として出されている現在の医師国家試験問題にも、同様の問題点があることが推測される。さらには、今後の医師国家試験の試験問題の作成に際して、コンテンツの専門家と試験問題作成法の専門家との緊密な連携体制を構築することが重要であることを示している。
- IT 相関が低い問題というのは、何か原因があるので、何が IT 相関の低さをもたらしているのかを検討することも、今後の研究班の研究課題となると思われる。

(2) 長期展望の点から今後の医師国家試験の体制構築にあたって検討すべきこと(表2)

- 多くの医療関係職種のみならず、教員採用試験など医療以外の分野でも、国家試験問題の作成にあたっては同じような問題を抱えている。
- 韓国では 26 の医療関連諸職種の国家試験を統括する Korea Health Personnel Licensing Examination Institute (KHPLEI) が置かれている。韓国では、医師国家試験は 1952 年にスタートし、1992 年までは韓国厚生労働省の管轄であったが、1992 年 4 月からは、KHPLEI (当初は Korea Health Personnel Licensing Examination Board という組織が設立された)が、徐々に対象の医療関係職種の数を増やして、現在では 26 の国家試験を管轄している。
- KHPLEI のような組織の構築は非常に重要であろう。

表2 各国の医師国家試験(相当試験)の CBT 化の比較

各国の医師国家試験 (相当試験) のCBT化の比較			
	韓国 (国家試験)	台湾 (国家試験)	米国(USMLE) (国家試験相当)
導入時期	2021	2015	1999
受験可能日	全国一斉(2日間)	全国一斉(2日間)	1年中
試験会場	全国9か所に19のCBT試験会場を確保。各種の国家試験に使用。施設は大学、国立の施設等	全国13か所 (会場は大学のコンピュータセンター)	Globally at a Prometric** centers
問題配布	USBで問題セットを各試験会場に持ち込んで実施	MOEの中央サーバーからInternetで配布	米国のサーバー (クラウド) から各国へInternetで配布
問題数	・320問 ・初年度にはマルチメディアを使った問題を3問出題 ・毎年3問ずつ増やす予定	・320問 ・初日 80問/2時間×3 2日目 80問/2時間×1	・280問(40問/1時間×7ブロック) ・休憩1時間はどこで取ってもよい ・同時に受験している受験生は、問題セット(280問)は同じで、出題順が異なる
合否判定主体	NHPLEI*	MOE†	NBME‡
安全性 (CBT vs CBT)	差なし	MOEは国家行政機関であり安全上の問題は感じていない	CBT化した時に、安全性を導入の理由の一つに挙げていた
試行期間	パイロット研究×2 試行1回	2年間 まず歯科医師国家試験から導入	・CBT化の決定: 1995 ・フィールドテストの開始: 1996 ・作問の準備: それまでの問題数の2倍の問題数を作問者に依頼 ・開始: 1996 ・CBTに移行: 1999
試験問題のプール	NO	NO	YES
IRTの導入	NO	NO	YES
医師以外の国家資格試験へのCBTの導入	医師を含め26の医療関連専門職の国家試験を統括している。	12種の医療専門職の資格試験がCBT化されている。2024にはNs, 栄養士の国家試験にも導入予定	対象は医師のみ

* National Health Personnel Licensing Examination Institute (Korea)

‡ National Board of Medical Examiners (USA)

** Test delivery company 試験実施会社

† Ministry of Examination (Taiwan)

3. 米国の USMLE (United States Medical Licensing Examination)について: 2020 年度の門田班における分担研究より抜粋(参考)

米国で医師免許を取得するにあたり、FSMB(州医療審議会連盟)とNBME(国立医療試験審議会)が実施主体となる USMLE の3段階の試験(STEP1~3)に合格することが必須となっている。(表3)

USMLE の STEP 1 は一般的に医学生 2 年時の終わりに受験し、基礎的な知識につい

て多肢選択問題(MCQ)で評価する。STEP 2 は 4 年次に受験し、医学知識、技能、臨床医学の理解を多肢選択問題(MCQ)で評価する。STEP 3 は卒後研修医 1 年目に受験し、総合的な医療知識、病態生理、臨床科学の理解について 2 日間の多肢選択(MCQ)とシミュレーションテストで評価する。

USMLE の大きな特徴は 1999 年より CBT を導入していることである。試験が年 2 回から一年中受験可能になり、STEP 1 と 2 の試験は一日に短縮された。試験は MCQ(多肢選択)形式。STEP3 のみ CCS(computer-

表3 米国医師国家試験(USMLE)の概要

	STEP1 (基礎医学分野)	STEP2 (臨床医学分野)	STEP3
評価内容	基礎的な知識を評価	医学知識、技能、臨床医学の理解を評価	総合的な医療知識、病態生理、臨床科学の理解を評価
対象	一般的な医学生においては2年時の終わりに受験する場合が多い	4年次に受験する場合が多い	卒後研修医1年目に受験する場合が多い
出題形式	1日の試験(MCQ) 解剖学、生理学、生化学、薬理学、病理学、行動科学	1日の試験(MCQ) 内科、外科、小児科、精神医学、産婦人科、公衆衛生、家庭医学、救急等	2日間の試験 1日目: 7時間(6ブロック(60分)に分割) ・MCQで256問 2日目: 9時間 ・MCQで198問(6ブロックに分割) ・ケース・シミュレーション(10~20分程度の13問)

based case simulation)形式のブロックが追加となった。受験は年中可能であり、STEP 1, 2 は各国のプロメトリックセンターで実施し、STEP 3 は米国内のプロメトリックセンターで実施する。

CBT 化の利点は、①セキュリティの向上、②試験形式の改良、③受験日の柔軟化、④効率的な試験運営が挙げられた。

筆記試験と CBT の比較では、成績の差は殆どなく、試験形態は成績に影響しないと結論づけられている。

USMLE では、試験ごと 1~300 点のスコアであり、IRT を利用し適応的に出題することで、受験回ごとのスコアは等化され比較可能としている(大きな改変の後には比不可)。また、60~70%の正答率で合格となるが、STEP3 については Case simulation が合否に大きく影響する。STEP 1 では、当初得点を発表していたが、現在は合否のみの発表となっている。

各 STEP において複数回受験は可能だが、同じ区分を 12 か月の間に 3 回までしか受験することができない。また、4 回目とそれ以降は最初の受験日から 12 か月以上かつ最後の受験から 6 か月以上あける必要があり、最大で 6 回までとなっている。すべての STEP は最初の STEP に合格した日から 7 年以内に合格しなければならない。

試験問題は、医療教育評価の専門家と臨床医からなる 2 つの試験委員会によって試験の材料・データを作成している。試験委員会のメンバーはアメリカの医療機関の委員会からなり、少なくとも 2 つの委員会は互いに試験問題や試行問題を批判的に鑑定し、疑問があれば改訂もしくは破棄する。試験問題は非公開だが、問題を集めている会社もある。

米国では USMLE を、臨床研修を行う病院とのマッチングや、自前の医師資格制度や教育制度を持たない国の子達を受け入れる際に役に立っている。

D. 考察

CBT 実施会社 4 社(プロメトリック、ピアソン VUE、教育測定研究所、CBTソリューションズ)の比較から、外資系企業と日本企業の違い

や会社の規模、セキュリティ面や費用、会場の規模、試験の実績については特色がさまざまであった。医師国家試験の CBT 化に向けて、試験会場の選定や手配、運用について現行の医師国家試験の実施状況を踏まえ、必要条件や課題を整理し、ベンダーについて検討する必要があると考える。

また、我が国の医師国家試験の運用や体制を検討するにあたり、全世界で様々な試験の CBT 化が進んでいる状況を把握するとともに、韓国の国家試験を統括する KHPLEI や、米国の USMLE などの事例を参考にすることが重要である。

また、CBT 化への準備過程においては、医師国家試験の受験方法について検討すべきである。すなわち、全国一斉試験ではなく、年間複数回の CBT 化医師国家試験を受験可能とし(ただし年間の受験回数の制限は行う)、臨床実習開始後は何時受験してもよいようにして、現在の 6 年生の臨床実習を卒業間際まで十分実施できるようにすべきである。

また長期的展望としては、医師国家試験を分野ごとにモジュール化(ex.3~6 分野)して、モジュールごとに受験を認め、最終的にはすべてモジュールに合格したら資格を与えるような受験方式も検討する余地がある。

E. 結論

医師国家試験の CBT 化に向けて、CBT 実施会社 4 社の比較をしたところ、試験会場の規模やセキュリティ面、同時実施人数の上限、会場費、運営にかかるコスト(人件費等)においてさまざまであったことから、引き続き検討する必要がある。

また、世界の国家試験の CBT 化の状況を踏まえて、試験の実施方法や受験方法について検討を要する。

さらには、医師以外の医療関連職種为国家試験の実施および組織体制の構築と運営などについて検討が必要である。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表
特になし
2. 学会発表
特になし

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
なし