

厚生労働科学研究費補助金
地域医療基盤開発推進研究事業

「医師国家試験の在り方に関する研究」

平成 27 年度 総括報告書

研究代表者 高木 康

平成 28 年 3 月

目次

- . 総括研究報告
 - 医師国家試験の在り方に関する研究 ----- 2

- . 会議議事録と参考文献
 - 1. 第1回会議
 - 1) 会議議事録
 - 2. 第2回会議
 - 1) 会議議事録
 - 2) 講演者スライド
 - 「自治医科大学と日本内科学会におけるコンピューターを使用したマルチメディア対応型試験の導入」 岡崎昭仁教授（自治医科大学）
 - 「CBT形式での新工夫 多数の画像の表示-画像のページング-」 青木茂樹教授（順天堂大学医学部）
 - 「台湾における医師国家試験-2015年7月18～20日視察-」 奈良信雄、高木 康（研究班員）
 - 3. 第3回会議
 - 1) 会議議事録
 - 2) 講演者スライド
 - 「アメリカ医師国家試験委員会（United States Medical Licensing Examination）」 奈良信雄、高木 康（研究班員）
 - 4. 第4回会議議事録
 - 1) 会議議事録
 - 5. 視察報告書
 - 「インドネシア国家試験 CBT 視察報告書」 大西弘高（東京大学医学部）

- . 参考書類
 - 厚生労働科学研究費補助金交付申請書
 - 倫理委員会承認書（昭和大学医学部倫理委員会）

医師国家試験の在り方に関する研究

研究代表者 高木 康（昭和大学医学部教授）

研究要旨

医師として国民から信頼される医療を実施するには、医療人としてのプロフェッショナルリズムを涵養し、十分な知識と技能を修得しており、医療人としての態度を涵養することが前提である。これら知識、技能および態度の修得の評価は、知識は必修項目、総論、各論から出題される MCQ500 題により行われている。

参加型臨床実習の student doctor（学生医）となるための共用試験での知識を問う問題は CBT（computer-based testing）により行われている。この CBT では通常が多肢選択肢問題（MCQ）だけではなく、後戻りできない機能を設定すること等により、医師として重要な臨床推論能力を評価する問題を作成することも可能である。

本研究では、医師国家試験の在り方としての知識を問う筆記試験の有効な手段としての CBT の現状と将来について調査研究を行った。すなわち、

諸外国での医師国家試験の筆記試験の実態を調査し、特に CBT 形式での出題形式、作成過程などを明確にする。

CBT 形式での利点と設備を含む問題点を明らかとして、医師国家試験に導入する際のシステムについて提言する。

ことを目的として、

台湾、米国での医師国家試験と CBT 導入の現状を視察して、医師国家試験の問題形式、作成過程、評価基準等の実態を調査した。

医師国家試験における CBT の活用の有効性、特にマルチメディア導入の現状と諸問題について、討議を行った。

調査を行った米国と台湾ではすでに CBT を医師国家試験に導入しているが、CBT への動画・マルチメディアの導入は計画はあるものの諸事情により行っていないことが確認された。また、この領域の専門家の講演では動画やマルチメディアの導入の可能性は高いが、我が国の医師国家試験の現状（多数受験生の一斉の受験など）では困難な点も多いが、動画やマルチメディアの導入による CBT は受験生のより深い知識の有無を評価できる可能性も高く、検討する価値が高いことが確認された。今後は小規模での実施を念頭にマルチメディアを使用した CBT の開発を行うことが我が国の医師国家試験の改善に繋がるものとの結論となった。

【研究組織】

研究代表者

高木 康（昭和大学医学部教授）

研究協力者

石田達樹（医療系大学間共用試験実施評価機構事業部長）

井廻道夫（新百合丘病院消化器・肝臓病研究所所長）

大西弘高（東京大学医学部講師）

鈴木利哉（新潟大学医学部教授）

奈良信雄（東京医科歯科大学特命教授）

野上康子（教育測定研究所研究開発部研究員）

A．研究目的

1. 医師国家試験の現状と改善

我が国の医師国家試験は、多肢選択肢問題（MCQ）形式500題（一般問題：200題、臨床問題：200題、必修問題：100題）により行われている。しかし、医師に必要な解釈や問題解決能力を評価するには問題作成に十分な吟味と工夫が必要であり、多大な知力と労力を費やしている。このような医師国家試験を改善して、解釈や問題解決能力を評価できる問題を出題することは国民の健康維持・増進に寄与できる医師を適切に選抜することにもなる。また、卒前の臨床実習が見学型の臨床実習から診療参加型実習（クリニカルクラークシップ）へと移行している。このクリニカルクラークシップでの学習成果を正確に評価するには改善されたシステムによる医師国家試験の開発が必要である。

本研究では、まず諸外国で導入されている医師国家試験のシステムの実態調査を行い、それらの現状と短所・長所を検討することにした。特にCBT（computer-based testing）については実施の有無と、実施している場合には問題形式、作成過程、評価基準等現地へ赴いて実施責任者・関係者に直接調査することとした。これにより諸外国での医師国家試験形式を明らかにし、我が国の医師国家試験回線のために導入可能は箇所について検討した。

2. CBTとその改善方法

医師には臨床推論的思考が必要であり、これを適切に評価できるシステムを医師国家試験にも導入する必要がある。この1つが諸外国でも導入され始め、我が国でも医療系大学間共用試験実施機構が医学生の前臨床実習前にその知識を評価する手段として導入しているCBTがある。このCBTでは、冊子方式の筆記試験には採用できない動画や音声などのマルチメディアを使用することでより深い・広い臨床推論に関する問題を出題できる可能性がある。

このCBTに関しては、諸外国ばかりでなく我が国の教養試験での実施状況を調査し、マルチメディアの使用が可能か、使用可能な場合の方式、問題点と有用性を検討することとした。まずは、医療系大学間共用試験で採用されているCBTの現状調査を行い、すでに画像を利用したCBTを実施している大学・学会を検索して、その現状と効用を直接調査するとともに、画像を

CBTに使用する際の問題点も聴取して検討することとした。

CBTの導入に関しては我が国の医学教育・医療現場との整合性を検証することが必要である。医学教育の内容（カリキュラム）に即した内容あるいは医学教育現場に即した内容であるべきである。これらを十分検証して、特に臨床推論や臨床判断能力を問う問題のありかたについては検討しなければならない。この点では、医療系大学間共用試験機構のCBTが良き指標となる。本システムへの応用を考えるべき1つの方法ではある。

医師国家試験の改善は、量的なこともあるが、質的な改善も必要であり、CBTの導入はその一助となると考える。本研究は、諸外国のCBTの現状（問題形式、作成過程、評価基準等）を調査して明らかとし、CBTシステム構築の長所・短所を検討し、これを我が国の医師国家試験に応用し、パイロット的な試験問題を作成して、検証することにある。

また、大学入試センター試験でもCBT方式の導入が検討されているなか、全国9,000人超の受験生への適応に向けての課題の抽出を併せて行う必要がある。

B．研究方法

1. 諸外国の医師国家試験のありかたの比較研究

アメリカ、カナダ、ドイツ、台湾、韓国などの医師国家試験が実施されている諸外国の試験制度で用いられている知識を評価する筆記試験、これに準ずる試験形式を調査し、我が国の医師国家試験の改善にあたっての参考にする。特にコンピュータを使用した試験システムとその内容について詳細な調査を行う。

2. コンピュータを用いた試験の国家試験への応用の検討

コンピュータを用いた試験（CBT）の現行の試験形式以外の領域への応用の検討を行う。現行のCBTでは冊子を用いた多肢選択肢問題（MCQ）の代用としての有用性により採用されている。CATOのCBTでは後戻りできない機能を追加して、臨床推論的な設問を出題している。その他に、現行では静止画像だけを出題しているが、患者の動作・動きやカラードップラなどを動画として設問する問題、心音や呼吸音を聴取して設問する問題など応用について検討する。特に、医師として必要な知識ばかりでなく技能や態度の評価への適応について検討する。

3. 医療系大学間共用試験実施機構（CATO）のCBTの内容の精査

現在、我が国のすべての医学部で臨床実習前にCATOの共用試験が行われており、技能態度についてはOSCEで行われ、知識についてはCBTにより評価されている。CATOで使用されているCBTシステムの詳細について検討し、医師国家試験に導入できるシステムを検討する。

4. 医師国家試験へのCBT導入に関する研究

現行の医師国家試験では冊子方式で500題が出題されているが、taxonomyでは想起・解釈領域の設問がほとんどである。また、CATOのCBTの良い点を再考察し、医師国家試験へ導入する場合の問題

点について検討するとともに、9,000人超が受験する医師国家試験への導入に向けた課題について抽出・検討を行う。

5. 医師国家試験のパイロットとしての試験問題の研究

CBT形式での医師国家試験の試作問題を作成し、試行してCBT形式の問題の問題点を明らかにして、改善のポイントを明確にする。

(倫理面への配慮)

医師国家試験の在り方を研究するもので、研究対象に倫理面での問題はない。

C. 研究結果

1. 諸外国での医師国家試験の現状

委員の奈良信雄、高木康研究員が台湾とアメリカの医師国家試験の、大西弘高研究員がインドネシアの医師国家試験の現状の現地調査を行った。

1) 台湾の医師国家試験

2015年7月17, 18日に台湾で行われた医師国家試験(CBT)を調査した。

医学部・受験生関連

- ・医学部数；13施設
- ・医学部学生：約1,300人(人数にばらつきがある)
- ・7年制(1～4年；講義、5～6年；クリニカルクラークシップ、7年；インターン) CBTのステップ(1)は4年次終了時に、インターン終了時にOSCEを行い、これに合格した者がステップ(2)を受験する。
- ・2013年入学生から6+2年制に移行した。1～4年は臨床実習前学習、5～6年はクリニカルクラークシップ、その後の2年間は臨床研修
- ・ステップ(1)受験は同じであるが、ステップ(2)はクリニカルクラークシップ終了時に行う。

CBT内容・実施

- ・ステップ(1)：解剖学・組織学・生理学・病理学などの基礎医学(70%)と臨床関連知識(30%)
- ・ステップ(2)：内科・外科・産婦人科など
- ・受験料：3,000元(約10,000円、OSEも同額)
- ・全国の医学部を4グループ(臺北・臺中・臺南・高雄)に分けてインターネットで各試験場に配信する。
- ・中央管理で行い、試験場のチーフ管理者(20年間勤務)がすべてのトラブルに対応する。

CBT試験問題

- ・出題問題数の10倍量(ステップ(1)：2,000題、ステップ(2)：3,200題)をプール化している。
- ・ブラッシュアップで各問題を「易」「中間」「難」に分類して30、40、30%の比率で出題するため、合格基準点は60%(絶対評価)としている。
- ・ブループリントとマニュアルは中国語で作成して、HP上にも公開している。
- ・マルチメディアは使用していないが考えている。
- ・試験問題はお暗示であるが、設問番号がラン

ダム化され、選択肢の順番も異なる。

- ・計算用紙を配布して、多少複雑な計算問題を出題している。
- ・開始後45分で退室可能である。

OSCE関連

- ・ステップ(2)の前にOSCEを行い、OSCEが合格しないとステップ(2)を受験できない。
- ・OSCEは5月に行う。
- ・1ステーション10分(実技8分、移動2分)で、12ステーションを受験する。
- ・8つがSPステーション、4つがクリニカルスキル(縫合など)である。
- ・合格率は約98%である。

2) アメリカでの医師国家試験

米国では医師国家試験(USMLE)はStep1、2、3があり、CS(Clinical Skill)以外にCBT(Computer-based testing)を導入している。

USMLEの歴史

- ・NBMEは1915年に発足して今年が100周年を迎える。米国・カナダ以外の外国医学部卒業生にはECFMGによる試験、米国人にはNBMEによる試験が行われていたが、これらがUSMLEとして1999年から統一された。
- ・CBTは1999年から行われた。これに先立ち1996～98年にCBT実施計画がなされ、1999年から実施された。CSは2004年にStep2に追加された。

CBT

- ・試験問題開発委員会が毎年新作問題を作成する。相当数の問題が作成される。ブラッシュアップされた新作問題は非採点問題として出題される。別の委員会により統計学的に許容された問題がプールされ、3年ぐらいで見直しが行われる。プール化に際しては日本と同様に200人以上に暴露して統計学的な有用性を検討する。IRT特性採用に関しては1パラメーターでは100名、2パラメーターでは200名が必要となる。
- ・一般的な問題作成法：作題者からNBMEへ問題が提出され、委員会でブラッシュアップ(BU)して、一部を作題者に返却して、昨題者による修正・改変が行われる。最終的には委員会で確認を行い、出題を決定する。
- ・新作問題は日本と同様にトライアルとして出題され、採点されない。評価委員会で統計学に基づいて検討され、良問はプールされ、修正した問題は再度トライアルされ、プール化され、あるいはダメな問題については削除される。プールされた問題は定期的にチェックされる。これは統計的数値の変動、新しい概念の導入およびガイドラインの変更により修正が必要とされるため、重要な作業である。日本の共用試験でもプール化委員会でプール問題を定期的にチェックしている。
- ・問題作成・出題の流れ：多領域点検委員会(IRCは内科や外科、産婦人科など多領域の先生が集まって新作問題の採択を決定する。偏った集団での評価は適切ではなく、難しい問題などが出題されるためである。期限切れになった問題はIRCで再評価して、修正を行う。良

問はIRCの指定した限りは3年間ぐらい使用する。

- CBTシステムの概要：受験者数は年間約30,000人が同じ会場で1年中受験している。毎年各ステップでのテストフォームが作成されている。30,000人のなかには外国人や繰り返し受験する受験者も含まれている。米国は約15,000~16,000人で、それ以外は外国人や繰り返し受験している。米国には145校の医学部があり、100~160人が在籍しているので、20,000人近くなるか

USMLE

- 運営規定：受験資格についてはStep1と2はLCMEかAOAで認証を受けた米国・カナダの医学部の学生または卒業生とECFMGによる資格認定を受けた米国・カナダ以外の医学部学生または卒業生である。Step 3はStep 1と2の合格者であるが、AOA卒業生も受験する。
- 運営方針：身体障害のある受験生への配慮があり、受験時間の延長や受験生への付き添い、音声記録などを行っている。受験回数は6回までである。受験間隔は1年間で3回までで、4回目受験は3回目受験から少なくとも6か月開ける。受験回数については、国家試験で検討されているが実施されていない。
- 合否判定：基準設定は、受験生のパフォーマンス(出来具合)、ステークホルダー(市長、市民代表、医学部学校の教員、看護師など)の意見、Angoff法、分析/分類誤差などを考慮して決定し、合格最低レベルは3~4年毎に見直す。ステークホルダーについては、LCMEでは学生もその中に入っている
- 受験数の推移：米国・カナダが80,000回、それ以外の外国が60,000回程度で、複数回受験もいて毎年30,000人程度が受験している。これらにはStep 1、2、3は入っているものと思われる。
- 結果報告：受験生に対しては、総得点と合否基準、分野別試験問題別の得点(良くできた分野、できなかった分野)が報告される。
- STEP 2 CKの報告書：STEP 2 CKについての概略が記載された後に、「合格もしくは不合格」と「点数」が記載され、さらに平均点とSD、SEMが記載されている。
- 結果報告：受験生の個人的な総得点、合格基準、分野別の得点シートが報告される。また、受験生からの要望があれば、免許交付の合格基準のために州の免許交付当局、進級/卒業判定での改善点のために医科大学、入学基準を知らせるために卒後プログラムへ情報が返却される。
- 管理システム：USMLEについては、ECFMG、FSMB、NBMEが親組織委員会、管理委員会、IRCs、TMDCs、その他の専門委員会などのすべての組織の管理に責任をもって運営している。
- 受験料：Step 1、2、3の受験料は、米国・カナダでStep 1、2が590\$、それ以外の学生では850\$で約1.5倍程度である。また、米国・カナダ以外の地域で受験する場合には地域により別途料金がかかる。

- USMLEに関する情報：USMLEのWebsiteで、試験に関する情報、試験サンプル、運営規則、試験に関するデータ、解釈のガイド、引用を閲覧できる。

3) インドネシアの医師国家試験

医学部関連

- カリキュラムは5年半(臨床前が3年半、臨床実習が2年間)
 - カリキュラムはOBE(2003年に国で制定)
 - 医学部数：75校
 - 医学生数：300人/年(ハサヌディン大学)
- #### 医師国家試験
- CBTとOSCEの両方の合格が必須
 - CBTとOSCEは4回/年(2、5、8、11月)
 - 受験料：OSCEが70万ルピア(約6,000円)、CBTが30万ルピア(2,600円)
- #### CBT
- 問題数：200問でAタイプ。全てが症例シナリオ問題
 - 受験時間は200分(1問1分)
 - 最大120名/回で、入学者数より少ない。
 - ブループリントと例題；開示している。
 - プール問題：各大学で作成した問題をプール化する。
 - 合格率：60%前後で、不合格者は3か月以降に再受験可能。
 - マルチメディア：画像は使用されている。
- #### OSCE
- ステーション：12St+レスト2Stの14Stで小外科、救急、小児科、産科、精神科は含まれる。
 - 医療面接+身体診察が主
 - テスト会場：大学に設置。
 - 課題：当日の開始1時間30分前に開封。
 - 評価者：1Stに1名。
 - SP：当日の1時間30分前に行い、外部のインパソン、芸術学部の俳優の卵などが動員される。時給は5万ルピア(約450円)。
 - 外部評価者：試験センター、近隣の大学から派遣され、全体視察、一部Stの評価を行う。
 - 合格ライン：80%で、1つでも不合格な場合は不合格。
 - 合格率：80%で、不合格の場合には3か月以降間隔を置いて再受験可能。

2. CBTのマルチメディア使用

CBTへのマルチメディア使用の可否について3名の経験を講演していただいた。

1) 「自治医科大学と日本内科学会におけるコンピュータを使用したマルチメディア対応型試験の導入」岡崎仁昭(自治医科大学)

医師育成過程の変化

- 医師国家試験：一般問題の減少とPCC-OSCEの導入
 - 内科学会資格認定試験：症候から優先順位を考慮しつつ鑑別診断を進める臨床医の思考過程に沿った臨床実地長文問題の導入
- #### 医師国家試験への提言
- 診療参加型臨床実習を真面目に行っている医学生に有利な内容に

- ・将来的にはmulti-media adapted CBT形式の導入
- ・医師国家試験が変わらなければ、診療参加型臨床実習（卒前教育）は機能しない。
コンピュータを使用したマルチメディア対応型試験の導入
- ・内科学会における専門医評価法
臨床実地長文連問形式問題の導入
- ・コンピュータを使用した試験問題
multi-media adapted CBT形式問題
- ・PCC-OSCE
臨床実地長文連問形式
- ・より実地臨床に即した臨床実地長文連問形式の導入
- ・実地臨床に即した新形式：医療面接、医療面接からの鑑別、身体診察、検査所見からの鑑別診断
- ・医師国家試験長文問題拡大形式
- ・multi-media adapted CBT:動画(手のふるえ、MMT)、聴診所見+心エコー図
自治医科大学でのM4総合判定試験
- ・Multi-media adapted CBT形式+解答は筆記式
- ・共用試験CBTとOSCEの相関は悪い年もあるが、multi-media連問と共用試験OSCEは強い相関が認められる。
- ・Multi-media連問とPCC-OSCEは相関を認める。
今後望まれる医師国家試験
- ・実地臨床に即した臨床実地長文連問形式の導入
- ・Multi-media adapted CBT形式の導入
- ・PCC-OSCEの医師国家試験への導入

2)「多数の画像の表示-画像のページング-」青木茂樹(順天堂大学医学部)

CBT形式での多数の画像の表示

- ・タイル表示
- ・Windowsのviewer
- ・Power Point
- ・DICOM
- ・クラウド利用の本格的なviewer
画像の特徴と解析法
- ・容量は多いようで、今となっては多いとは言えない。
タイル表示
- ・現状のソフトですぐに対応可能である。
- ・画像が小さいので、解像度が問題である。
- ・下図を限った表示（12枚程度まで）なら可能か。
ページング表示
- ・タイル表示よりページング表示の方が視点を動かさずに見ることができ、観察が可能である。
- ・タイル表示より解像度は良好となる。
- ・PACSでの観察（現状の多くの施設で行われている）により近い。
- ・Windows付属の基本ソフトで対応可能か。
- ・現状のソフトの変更が必要（これ以下の項目に共通する問題）。
- ・T1強調像、T2強調像、FLAIR、造影前後、肺野条件・縦隔条件など多くの画像を見るには

適していなし。条件の変更もできない。

DICOM viewer

- ・実際であるがソフト・データ集積に問題があるか。
クラウド利用
- ・実際的でIT関連の進歩を考えるとbetterな方法か。
- ・データの蓄積、管理を含めたシステムとなり、後の管理は容易なはず。
- ・ただし、外部との通信が必要であるのが、欠点である。
- ・個人情報保護法はクリアしている？データ登録の際に匿名化が必要。

3)「米国医師国家試験CCSの紹介」片岡仁美(岡山大学医学部)

CCS: computer-based case simulation

ステップ3: 467問

内容

- ・時系列あるいは別の状態を設定している。
- ・救急症例、通常の外来症例などを対象とする。
回答時間は症例による。
- ・症例に対する対応・処置を順次質問してくるので、それに対して対応に関して解答する。
- ・フロー的な質問と解答を組み合わせる。
- ・依頼の項目は解答者が考え、コンピュータに入力するので、難しい。
- ・検査とマネジメントを組み合わせることが大切。症例対応・対処が重要である。
- ・PBL形式となっている。
- ・処置を変えることはできる。
- ・米国では6か所に分散して、試験を行っている。
- ・Common diseaseが多い。
- ・13症例には緊急症例も多い。適切な対応・処置が必要である。

D. 考察

本研究では、外国（台湾、アメリカ、インドネシア）の医師国家試験について視察し、CBTを含めた現状と問題点を検討した。今回、視察した国では医師国家試験としてCBTを導入している、マルチメディアについては、画像は使用されているが、動画や音声などは使用されていない。アメリカでは、音声については一部使用されているが、動画については現在検討中とのことである。

我が国では医療系大学間共用試験実施評価機構が主催する共用試験でCBTが実施されている。このCBTをこれら3か国のCBTと比較すると、概要はほぼ同等な内容である。共用試験CBTでも動画や音声などのマルチメディアの使用が検討されている。深い知識を評価するにはマルチメディアを使用したCBTもその有効な手段の1つと考えられる。

E. 結論

我が国の医師国家試験は冊子による筆記試験が行われている。この様式では、深い知識を評価するのは必ずしも容易ではなく、補完する手段を考える必要がある。マルチメディアを使用したCBTはその1つの手段であり、今後はこれを利用したCBTの開発が必要である。また、視察した3か国では

医師国家試験としてCBTとOSCEが併用されており、両方を合格することが必須となっている。我が国でも医科大学・医学部によるPCC=OSCE (post-clinical clerkship OSCE)の実施が計画されており、マルチメディアを使用したCBTとOSCEにより国民の健康増進に貢献する医師の育成が行われることが期待される。

F．健康危険情報

該当なし。

G．研究発表

- 1.論文発表
国家試験に係ることで、論文の公表はなし。
- 2.学会発表
国家試験に係ることで、学会での公表はなし。

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

【謝辞】

本研究の遂行にあたっては、厚生労働省医政局医事課試験免許室のご協力をいただいた。ここに深謝する。

G . 研究発表

1. 論文発表
国家試験に係ることで、論文の公表はなし。
2. 学会発表
国家試験に係ることで、学会での公表はなし。

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

第 1 回研究会議 議事録

1. 日時：平成 27 年 6 月 23 日（火） 15：00～17：00
2. 場所：公益社団法人医療系大学間共用試験実施評価機構会議室
（文京区湯島 1-9-15 御茶ノ水 HY ビル）
3. 参加者：石田達樹、井廻道夫、大西弘高、鈴木利哉、奈良信雄、野上康子、高木 康
4. 議題

1) 自己紹介

2) 厚労科研申請書の説明（高木）

前厚労省医政局医事課試験免許室試験専門官大淵氏の構想に基づく委託研究。成果は次回の医師国家試験改善検討部会に利活用される（と関連）可能性諸外国での医師国家試験の筆記試験での実施を調査し、特に CBT 形式での出題形式、作成過程などを明確にする。

- ・ 諸外国での CBT の実情調査
- ・ 諸外国医師国家試験の CBT でのマルチメディア活用法の調査

CBT 形式での利点と設備を含む問題点を明らかにして、医師国家試験に導入する際のシステムについて提言する。

- ・ 現在 CBT 形式で出題されている試験の現状調査
- ・ CBT 形式に利用できるマルチメディアの活用法

マルチメディア・・・文字や画像、動画、音声など、様々な種類・形式の情報を組み合わせて複合的に扱うこと、さらには、利用者の操作に応じて表示や再生の仕方に変化が生まれる双方向性も含む（クリックマップ等）

3) 本研究での確認事項

国家試験の筆記試験への CBT 適用

- ・ 試験回数---年 1 回（現行を踏襲）
- ・ 現行筆記試験のどの程度を置換えるか---全部/一部
- ・ CBT 適用時の課題---試験センターの確保、機器の整備、機器使用時の不具合対応など運用面の問題が存在（機器関連の課題は一斉実施に関連する問題であり、現行の共用試験のような逐次実施では問題になりにくいという違いがある点に注意）
- ・ 厚労省が試験を実施する場合は試験実施後の試験問題漏洩・再現は問題とならないが、問題プール制を前提とした場合には受験生間の公平性の観点から十分な注意を要する。

CBT 利用の利点

- ・ 視覚障害者への対応可能
- ・ 多数の画像（例えば CT のマルチスライス）、動画や音声などを利用した試験が可能
- ・ ペーパーベースでは実現が困難な問題出題形式の利用（例えば、1 度解答して次の問題に進むと後戻りできない形式）

4) 研究検討事項の討議

医学領域や他領域での CBT の現状

- ・ 医学ばかりでなく、歯学・薬学でも CBT を利用しており、看護や獣医学領域（平成 28 年度から正規実施予定）でも CBT 実施の予定である。
- ・ 医療系領域ばかりでなく、他の領域の資格試験などでの利用の内容などの調査を行う。

諸外国での国家試験での CBT 形式の利用調査

- ・ 北米や欧州、韓国での国家試験（OSCE）については、すでにかなり調査が行われているが、CBT 形式についての調査は少ない。
- ・ 北米やアジア諸国の実情を調査する。
- ・ 韓国は未だ利用されていないが、現状を調査する。
- ・ 台湾は今年度から実施する（？）予定であり、現状を調査する。
- ・ その他の国の CBT 形式の国家試験の内容を調査する。

コンピュータを利用した試験

- ・ 音声や動画を利用した試験が実施できる。
- ・ 電子カルテを模した画面参照（例えば、検査結果情報を問題文に埋め込むのではなく、検査結果参照 WINDOW を表示して見せるなど）
- ・ 画像の中に色々なリンク先を設定しておき、クリックした位置に応じて定められたリンク先に移動するクリッカブルマップ機能やイメージ上の特定範囲をマウスで選択する機能の利用

臨床推論（共用試験 CBT ではブロック 6 で出題）への利用

- ・ 順次解答形式連問による試験の有用性を検討する。
- ・ 臨床推論試験としては優れているので、内容について吟味する。

海外訪問しての国家試験実情調査

事前に十分な下調べ等を行う

- ・ 米国 奈良先生、大西先生、高木先生
- ・ 台湾 奈良先生、大西先生、高木先生
- ・ 韓国 大西先生（計画も含めて CBT の実情）、高木先生

5) 今後の予定と分担

諸外国（北米、アジア）の国家試験関連団体に現状と訪問の都合を質問する。
我が国のマルチメディア機能を用いた出題形式の先行事例について把握する。

- 1回目 - （他の先行事例があれば追加で行う）

・青木茂樹先生（順天堂大学） 画像のページング

・岡崎仁昭先生（自治医科大学） 内科専門医試験へのマルチメディアの活用
両氏には次回の班会議で CBT 形式の各種試験への応用について、講演していただく。

日程調整は8月11日、13日、14日を候補日として高木研究班長が行う。

なお、第2回以降の研究班会議から大淵専門官の後任の鈴木専門官に対して可能であれば陪席を要請する。

次回会議

日時：平成27年8月13日（木） 10:00～12:00

場所：公益社団法人医療系大学間共用試験実施評価機構6階会議室

講師：青木茂樹（順天堂大学）、岡崎仁昭（自治医科大学）

平成 27 年度厚生労働科学研究費 「医師国家試験の在り方に関する研究」

日時:平成 27 年 8 月 13 日(木) 10:00~13:00

場所:公益社団法人 医療系大学間共用試験実施評価機構

6 階会議室 東京都文京区湯島 1-9-15 茶州ビル

参加者:石田達樹、大西弘高、鈴木利哉、奈良信雄、野上康子、高木 康

アドバイザー:青木茂樹、岡崎仁昭、片岡仁美

議題

1. CBT への導入手法の講演

1) 岡崎仁昭先生

題名:「自治医科大学と日本内科学会におけるコンピュータを使用したマルチメディア対応型私権の導入」

医師育成過程の変化

- ・ 医師国家試験:一般問題の減少と PCC-OSCE の導入
- ・ 内科学会資格認定試験:症候から優先順位を考慮しつつ鑑別診断を進める臨床医の思考過程に沿った臨床実地長文問題の導入

医師国家試験への提言

- ・ 診療参加型臨床実習を真面目に行っている医学生に有利な内容に
- ・ 将来的には multi-media adapted CBT 形式の導入
- ・ 医師国家試験が変わらなければ、診療参加型臨床実習(卒前教育)は機能しない

コンピュータを使用したマルチメディア対応型試験の導入

- ・ 内科学会における専門医評価法
臨床実地長文連問形式問題の導入
- ・ コンピュータを使用した試験問題
multi-media adapted CBT 形式問題

- ・ PCC-OSCE

臨床実地長文連問形式

- ・ より実地臨床に即した臨床実地長文連問形式の導入
- ・ 実地臨床に即した新形式:医療面接、医療面接からの鑑別、身体診察、検査所見からの鑑別診断
- ・ 医師国家試験長文問題拡大形式
- ・ multi-media adapted CBT:動画(手のふるえ、MMT)、聴診所見+心エコー
図

自治医科大学での M4 総合判定試験

- ・ Multi-media adapted CBT 形式 + 解答は筆記式
 - ・ 共用試験 CBT と OSCE の相関は悪い年もあるが、multi-media 連問と共用試験 OSCE は強い相関が認められる。
 - ・ Multi-media 連問と PCC-OSCE は相関を認める。
- 今後望まれる医師国家試験
- ・ 実地臨床に即した臨床実地長文連問形式の導入
 - ・ Multi-media adapted CBT 形式の導入
 - ・ PCC-OSCE の医師国家試験への導入
- 新医師国家試験 400 題

2) 青木茂樹先生

題名:「多数の画像の表示-画像のページング-」

CBT 形式での多数の画像の表示

- ・ タイル表示
- ・ Windows の viewer
- ・ Power Point
- ・ DICOM
- ・ クラウド利用の本格的な viewer

画像の特徴と解析法

- ・ 容量は多いようで、今となっては多いとは言えない。

タイル表示

- ・ 現状のソフトですぐに対応可能である。
- ・ 画像が小さいので、解像度が問題である。
- ・ 下図を限った表示(12 枚程度まで)なら可能か。

ページング表示

- ・ タイル表示よりページング表示の方が視点を動かさずに見ることができ、観察が可能である。
- ・ タイル表示より解像度は良好となる。
- ・ PACS での観察(現状の多くの施設で行われている)により近い。
- ・ Windows 附属の基本ソフトで対応可能か。
- ・ 現状のソフトの変更が必要(これ以下の項目に共通する問題)。
- ・ T1 強調像、T2 強調像、FLAIR、造影前後、肺野条件・縦隔条件など多くの画像を見るには適していなし。条件の変更もできない。

DICOM viewer

- ・ 実際的であるがソフト・データ集積に問題があるか。

クラウド利用

- ・ 実際的で IT 関連の進歩を考えると better な方法か。
- ・ データの蓄積、管理を含めたシステムとなり、後の管理は容易なはず。
- ・ ただし、外部との通信が必要であるのが、欠点である。
- ・ 個人情報保護法はクリアーしている？データ登録の際に匿名化が必要。

3) 片岡仁美先生

題名:「米国医師国家試験 CCS の紹介」

CCS: computer-based case simulation

ステップ3: 467 問

内容

- ・ 時系列あるいは別の状態を設定している。
- ・ 救急症例、通常の外来症例などを対象とする。回答時間は症例による。
- ・ 症例に対する対応・処置を順次質問してくるので、それに対して対応に関して解答する。
- ・ フロー的な質問と解答を組み合わせる。
- ・ 依頼の項目は解答者が考えるので、難しい。
- ・ 検査とマネージメントを組み合わせることが大切。症例対応・対処が重要である。
- ・ PBL 形式となっている。
- ・ 処置を変えることはできる。
- ・ 米国では6箇所分散して、試験を行っている。
- ・ Common disease が多い。
- ・ 13 症例には緊急症例も多い。適切な対応・処置が必要である。

2. 台湾(中華民国)の医師国家試験 CBT 試験

別紙参照

3. その他

次回の開催は 11 月の奈良教授、高木の米国施設後に計画する。9 月中旬に予定を調査する。

台湾における医師国家試験 -2015年7月18~20日視察-



奈良信雄 (順天堂大学/東京医科歯科大学)
高木 康 (昭和大学)

1

台湾の医学教育概略

- 高校卒業生を教育
- 医学部は13校
- 6年制+2年 (1~4年: 臨床実習前学習、5、6年: クリニカルクラークシップ、+7、8年: 臨床研修)
- ~2012年までは7年制 (1~4年: 講義、5、6年: クリニカルクラークシップ、7年: インターン)
- 6月に学期が終了。卒業後1か月以内 (7月) に国家試験を受験
- 2015年の国試受験はStep1が2,211名、Step2が1,322名

2

台湾の医師国家試験運営

- 考試院が実施 (Ministry of Examinations)
- MOE x の組織：約200名のスタッフ
- 試験の取り扱い：医師、看護師、検査技師、工学系などの専門職試験 (約30万人受験) + 公務員、警察官等の試験 (約30万人受験)



医師国家試験概略

-Step 1-

- 4年生終了時に受験 (7月)
- 内容
 - ✓ 基礎医学 (解剖学、発生学、組織学、微生物学、寄生虫学、公衆衛生学) および関連した臨床知識：100題/120分
 - ✓ 基礎医学 (生理学、生化学、薬理学、病理学) および関連した臨床知識：100題/120分
- 合格率：約62%!!
- 不合格者は翌1月にチャンス

医師国家試験概略

-Step 2-

- 6年生卒業後1か月以内に受験（7月）
- 内容
 - ✓ 内科・家庭医学（知識・症例・倫理）：80題/120分
 - ✓ 小児科・皮膚科・神経内科・精神科（知識・症例・倫理）：80題/120分
 - ✓ 外科・整形外科・泌尿器科（知識・症例・倫理）：80題/120分
 - ✓ 麻酔科・眼科・耳鼻科・産婦人科・再生医療（知識・症例・倫理）：80題/120分
- 合格率：約94%
- 不合格者は次年度にチャンス（再受験の回数制限はない：長年受験者はまれ）
- 参考：OSCE合格率 約98%
- 国試ではCBT、OSCE両者の合格が必要。どちらかが不合格の場合には不合格科目のみ再受験

国家試験実施



- Step 1：7月と1月
- Step 2：7月（CBTとOSCE両者に合格が必須）
- CBT
 - ✓ 2004年：航海士の試験が台北試験会場で開始。
 - ✓ 2007年：歯科医、助産婦、呼吸治療師（台湾特異？）、獣医師試験
 - ✓ その後、歯科、薬剤師、検査技師、作業療法士、理学療法士、呼吸治療師、獣医
 - ✓ 2015年：医師
 - ✓ 台北、台中、台南、高雄の試験センターで実施。
 - ✓ 遅刻
 - 1時限：15分遅刻可、2時限以降：3分遅刻可
 - ✓ 退出
 - 試験開始後45分は退出不可

CBTシステム

- 試験会場：台湾国内4カ所（台北、台中、台南、高雄）
- 全国の試験実施状況は台北のセンターコントロールルームで確認
- 試験当日・試験直前に試験問題をインターネットで各試験場に送信
- パスワードで問題を開く
- セキュリティに配慮
- PCの予備も万全、オンラインで練習可能
- OSはWindows
- 受験生の解答はサーバーにリアルタイムでバックアップ
- 不測の事態に備える。無停電供給有り
- 成績は終了後に即採点（受験生は確認可）

7

CBT試験問題関連

- 4肢選択形式
- 基本的に約10倍の問題をプール（Step 1：2,000題、Step 2：3,200題）
- 各学会の推薦委員が問題を作成
- ブラッシュアップ（主文や選択肢は記載されている教科書のページと行目を記載）
- ブラッシュアップで問題を「易」「中間」「難」に分類して、30、40、30%の比率で出題
- 難易度を調整することで、合格基準を60%にできる
- IRTは使用していない
- カラー静止画像は出題されるが、動画、音声の出題はない（今後開発の予定）
- 画面上の画像の拡大機能あり
- 試験問題は統一しているが試験番号と選択肢をランダム化することで異なる問題としている

8

OSCE関連

- Step 2のCBTの前に行い、OSCEが合格しないと、Step 2を受験不可
- OSCEは5月に実施
- ステーション数は12（8つはSPステーション、4つはクリニカルスキル（縫合など））
- 1ステーションは10分（実技：8分、移動：2分）
- 全国に23センター（大学と大規模病院）で行い、評価者は厚労省所属（Minister of Health and Welfare）
- 5～8年後にOSCEセンターを考試場内に建設予定（13階建て、5階がOSCE専用フロア。12のステーションルームを予定）
- 4区間のOSCEセンターはそのままにして、OSCEセンターで集約的にOSCEを施行予定
- 今年度の合格率は約98%。

9



中華民國國家試驗 CBT 施設報告書 2015

視察日時：2015 年 7 月 17～20 日

場所：台湾考選部 (examination center)

面会者：周玉山 (考試院考試委員長)、高鳳仙 (監察委員)、董保城 (考選部部長)、
黃慶章 (考選部司長)、劉克明 (高雄医学大学教授)

視察者：奈良信雄、高木 康

1. 医学部関連

- 1) 学年
- 2) 医学部数；13 施設
- 3) 医学部学生：約 1,300 人 (人数にばらつきがある)
- 4) 7 年制 (1～4 年；講義、5～6 年；クリニカルクラークシップ、7 年；インターン) CBT のステップ(1)は 4 年次終了時に、インターン終了時に OSCE を行い、これに合格した者がステップ(2)を受験する。
- 5) 2013 年入学生から 6 + 2 年制に移行した。1～4 年は臨床実習前学習、5～6 年はクリニカルクラークシップ、その後 2 年間の臨床研修 ステップ(1)受験は同じであるが、ステップ(2)はクリニカルクラークシップ終了時に行う。

2. CBT 関連

- 1) ステップ 1：医学(1)と(2)；解剖学・組織学・生理学・病理学などの基礎医学 (70%) と臨床関連知識 (30%)
- 2) ステップ 2：内科・外科・産婦人科など)
- 3) 日程：7 月 18 日

ステップ(2)

9:00～11:00：医学(01)包括内科・家庭医学 80 題/120 分

12:55～14:55：医学(02)包括小児科・皮膚科・神経科・精神科 80 題/120 分

15:20～17:20：医学(03)包括外科・骨科・泌尿器科 80 題/120 分。

7 月 19 日

9:00～11:00：医学(04)麻酔科・眼科・耳鼻咽喉科・産婦人科・復健科 80 題/120 分。

受験生：1,322 名で、合格率は 90%程度。今年度は約 94%。

不合格者：次年度にチャンスがあり、再試験の回数に制限はない。

ステップ(1)

12:55～13:55：医学(01) 解剖学・組織学など 100 題/120 分。

15:29～17:20：医学(02) 生理学・生化学・病理学など 100 題/120 分。

受験生：2,213 名で合格率は 50%程度。台湾以外の医学部を卒業した受験生も

ある。 試験終了直後に合格率が出て、今年は約 62%。

受験料：3,000 元（OSCE の同額）

- ・ 全国の医学部を 4 グループ（臺北・臺中・臺南・高雄）に分けて、一斉に実施している。
- ・ 試験当日・試験直前に試験問題をインターネットで各試験場に送信する。パスワードで開く。
- ・ すべては中央管理で試験場のチーフ管理者（20 年間勤務）がすべてのトラブルに対応するシステムである。
- ・ 各グループの管理者は中央から派遣されており、チーフ管理者にトラブル内容を知らせて、適切な対応を指令される。
- ・ トラブルがあった場合には、職員が問題を輸送する。

4) 問題のプール化に関して

ステップ(1)と(2)：基本的には約 10 倍の問題のプール化を行っている。

ステップ(1)：2,000 題、ステップ(2)：3,200 題

各学会の推薦委員が問題を作成する。

ブラッシュアップを行う。主文や選択肢については、記載されている教科書のページと行目を記載する。

ブラッシュアップでは「易」「中間」「難」に分類して 30、40、30%の比率で出題する。合格点を 60%とするための方策。

合格点は 60%としている。絶対評価であり、IRT は用いていない。

5) 試験問題

公開している。

このため、問題の復元は行われていない。

医系 CBT は今年度が初めてであるので、今後のことは分からない。

6) ブループリントなど

ブループリントは作成している（中国語）

マニュアルも作成している（中国語）

これらは HP に提示されている。

7) CBT での工夫

マルチメディアは使用していないが、現在使用を考えている。

（歯学系では動画がある？）

画面上の写真を拡大することができる。

試験問題は統一されているが、試験番号はランダム化され、選択肢の順番も異なっている。

OS は Windows である。

45分終了時に退出できる。これは歯学系での60分試験でも医学系の120分試験でも同様である。

計算紙を配布して、多少複雑な計算問題を出題している。

8) CBTの導入

医科が2015年だが、2004年に歯科・薬科で実施されている。

コンピュータは会社のものを使用している(台湾製)

3. 全般

受験料：OSCEとCBTともに3,000円で合計6,000円

受験は2回/年で、7月と2月の2回。

受験票：写真は添付されていない。大学が発光しているIDを携帯するので、これを用いて本人確認を行っている。

CBTの年間スケジュールは1年前、7月をめどに行っている。

コンピュータ専門職が5名、会場に待機して備えている。

CBTとOSCEの合格が必要であり、不合格の場合には不合格科目のみ再受験。

4. OSCE

ステップ(2)のCBTの前に行い、OSCEが合格しないと、ステップ(2)を受験できない。

OSCEは5月に行う。

ステーション数は12で、8つはSPステーション、4つはクリニカルスキル(縫合など)である。

1ステーションは10分(実技：8分、移動：2分)

全国に23センター(大学と大規模病院)で行い、評価者は厚労省所属?(Minister of Health and Welfare)

5～8年後にOSCEセンターを考試場内に建設予定であり、13階建てでえ、5階がOSCE専用フロア。12のステーションルームを構築予定である。

この時には、4区間のOSCEセンターはそのままにして、1つだけ残し、OSCEセンターで集約的にOSCEを施行予定である。

合格率は約98%。

平成 27 年度厚生労働科学研究費 「医師国家試験の在り方に関する研究」
第 3 回会議議事録(案)

日時:平成 27 年 12 月 11 日(木) 15:00 ~ 17:00

場所:公益社団法人 医療系大学間共用試験実施評価機構

6 階会議室 東京都文京区湯島 1-9-15 茶州ビル

参加者:石田達樹、大西弘高、鈴木利哉、奈良信雄、高木 康

アドバイザー:鈴木貴士専門官

議題

1. 米国視察の報告(奈良委員、高木委員)

1) USMLE (United States Medical Licensing Examination)

議題:米国では医師国家試験(USMLE)は Step1、2、3 があり、CS(Clinical Skill)以外に CBT(Computer-based testing)を導入している。

USLME の歴史:NBME は 1915 年に発足して今年が 100 周年を迎える。米国・カナダ以外の外国人には ECFMG が行われている。米国人には NBME が行われており、これらが統一されて USMLE として、1999 年から統一された。

USLME の CBT 実施経時的推移:1999 年から行われた。これに先立ち 1996 ~ 98 年に CBT 実施計画がなされ、1999 年から実施された。CS は 2004 年に Step2 に追加された。

CBT の長所:CBT の長所としては、試験紙と鉛筆による試験より機密保持が優れている。試験紙による試験ではカンニングがあり、また米国は広いので試験問題用紙が配送中に盗難に会うことも考えられる。また、シミュレーション、ビデオ、音声を取り入れた試験も可能である。

CBT の短所:一方、CBT の短所としては、大量の問題をプールしなければならない。米国では年間約 30,000 人が受験するため、質の揃った良問をプール化する必要がある。試験紙問題と比較して年間いつでも実施できるが、そのための多数のスタッフが必要となる。

CBT 試験システム:試験問題開発委員会が毎年新作問題を作成する。相当数の問題が作成される。ブラッシュアップされた新作問題は非採点問題として出題される。別の委員会により統計学的に許容された問題がプールされ、3 年ぐらいで見直しが行われる。プール化に際しては日本と同様に 200 人以上に暴露して統計学的な有用性を検討する。IRT 特性採用に関しては 1 パラメーターでは 100 名、2 パラメーターでは 200 名が必要となる。

一般的な問題作成法: 作題者から NBME へ問題が提出され、委員会でブラッシュアップ (BU) して、一部を作題者に返却して、昨題者による修正・改変が行われる。最終的には委員会で確認を行い、出題を決定する。

新作問題の扱い: 新作問題は日本と同様にトライアルとして出題され、採点されない。評価委員会で統計学に基づいて検討され、良問はプールされ、修正した問題は再度トライアルされ、プール化され、あるいはダメな問題については削除される。プールされた問題は定期的にチェックされる。これは統計的数値の変動、新しい概念の導入およびガイドラインの変更により修正が必要とされるため、重要な作業である。日本でもプール化委員会でプール問題を定期的にチェックしている。

問題作成・出題の流れ: 多領域点検委員会 (IRC; Interdisciplinary Review Committees) は内科や外科、産婦人科など多領域の先生が集まって新作問題の採択を決定する。偏った集団での評価は適切ではなく、難しい問題などが出題されるためである。期限切れになった問題は IRC で再評価して、修正を行う。良問は IRC の指定した限りは 3 年間ぐらい使用する。

CBT システムの概要: 受験者数は年間約 30,000 人が同じ会場で 1 年中受験している。毎年各ステップでのテストフォームが作成されている。30,000 人のなかには外国人や繰り返し受験する受験者も含まれている。米国は約 15,000 ~ 16,000 人で、それ以外は外国人や繰り返し受験している。米国には 145 校の医学部があり、100 ~ 160 人が在籍しているため、20,000 人近くになるか。

USMLE 運営規定: 受験資格については Step1 と 2 は LCME か AOA で認証を受けた米国・カナダの医学部の学生または卒業生と ECFMG による資格認定を受けた米国・カナダ以外の医学部学生または卒業生である。Step 3 は Step 1 と 2 の合格者であるが、整骨医学校卒業生も受験する。

USLME の運営方針: 身体障害のある受験生への配慮があり、受験時間の延長や受験生への付き添い、音声記録などを行っている。受験回数は 6 回までである。受験間隔は 1 年間で 3 回までで、4 回目受験は 3 回目受験から少なくとも 6 か月開ける。受験回数については、国家試験で検討されているが実施されていない。

試験問題配信ソフト: 試験問題配信ソフトは NBME が開発している。SE は 200 人程度で、著作権も NBME にある。現在は、動画は入っていない。音声については CS では導入されているが CK では採用されていない。SE の 200 人は正規ではなく、パートやボランティアも含めてである。10 数年前は私的会社が作成していた。NBME の職員は全部で 400 人程度であり、法律家そ

の他の多くの専門家が職員として働いている。心音(聴診)はCSに導入されているが、動画については今後の課題である。

合格判定基準:基準判定は、受験生のパフォーマンス(出来具合)、ステークホルダー(市長、市民代表、医学部学校の教員、看護師など)の意見、Angoff法、分析/分類誤差などを考慮して決定し、合格最低レベルは3~4年毎に見直す。ステークホルダーについては、LCMEでは学生もその中に入っている。

USLMEの受験数の推移:米国・カナダが80,000回、それ以外の外国が60,000回程度で、複数回受験もいて毎年30,000人程度が受験している。これらにはStep 1、2、3は入っているものと思われる。

初回受験者の合格率:米国・カナダの受験者の合格率は高く、それ以外の国の受験者の合格率は低いが、徐々に上昇してきている。

結果報告:受験生に対しては、総得点と合否基準、分野別試験問題別の得点(良くできた分野、できなかった分野)が報告される。

STEP 2 CKの報告書例:結果報告書例で、STEP 2 CKについての概略が記載された後に、「合格」で「点数 199」と記載されている。平均点とSDは229と23であり、受験者の得点は140~260にあり、合格基準点は189点で、SEM(Standard error of measurement)は7ポイントと説明がある。

各分野の得点:各分野別の得点が記載されているパフォーマンスシートである。< Physician task profile > < Normal condition & disease category profile > < Discipline profile > にそれぞれがどの程度の得点かを提示してある。

結果報告:前記のような受験生の個人的な総得点、合格基準、分野別の得点シートが報告される。また、受験生からの要望があれば、免許交付の合格基準のために州の免許交付当局、進級/卒業判定での改善点のために医科大学、入学基準を知らせるために卒業プログラムへ情報が返却される。

- 21 結果報告 2:全受験者のデータ(総数と合格率、試験項目・分野別の得点・プロファイル)が報告される。
- 22 STEP 1のデータ:2014年の最初のSTEP 1のデータである。学生の所属する医学部の受験者数、米国・カナダ人の受験者について、初回と繰り返し受験の受験者数、合格者数、合格率が記載され、平均値とSDも記載されている。また、説明文には合格基準が192と記載されている。
- 23 総得点の分布:2014年の初回受験者の総得点の分布で、その医学部と全国の分布が報告される。
- 24 STEP1の分野別得点:医学部の初回受験者の各分野別(Behavioral Science から Biostatistics & Epidemiology/Population Health)について全国の

平均値と1SD(-1と1)と同時にその医学部の平均値(赤)と1SDが示されている。

- 25 USMLEの管理システム:USLMEについては、ECFMG、FSMB、NBMEが親組織委員会(Composite Committee)、管理委員会(Management Committee)、IRCs、TMDCs(試験教材開発委員会)、その他の専門委員会などのすべての組織の管理に責任をもって運営している。
- 26 受験料:Step 1、2、3の受験料は表のごとくであり、米国・カナダでStep 1、2が590\$、それ以外の学生では850\$で約1.5倍程度である、また、米国・カナダ以外の地域で受験する場合には地域により別途料金がかかる。
- 27 USLMEに関する情報:USLMEのWebsiteで、試験に関する情報、試験サンプル、運営規則、試験に関するデータ、解釈のガイド、引用を閲覧できる。

2) 奈良委員補足説明

米国における医学教育と医師国家試験制度 (USMLE & ECFMG)

改善を要する緊急課題

72週は州で決めている。外形基準である。3年で48週、4年で40~42週の臨床実習を行っている。臨床実習の間に知識を修得する。

USLMEの概略:Step2はCK(知識)とCS(技術)を試験し、指導者の下で医療を実践できる能力の評価であり、Step3が合格すると単独で医療を行うことができる知識とその応用を評価する。NBMEとFSMBが行っている。

Step 2 CS:7時間/日で、12ステーションで行う。SPを対象に医療面接、身体診察を行うが、電話での試験がある。小児例での母親役への説明と説得を評価する。SPが評価する。

ECFMG見学:米国・カナダ人と国際人が6人ずつで受験していた。試験は録画されている。

3) CBTの新展開・今後

動画を導入するが容量が大きくて難しいが、tremorなどの診断をするには導入する必要がある。

画像の導入が可能か。国家試験では、キーフィルムしか提示されていないので、全ての画像から受験生が異常を検出する必要がある。

音声は導入可能か。心音を聴取させるが、胸部の部位を受験生が指定して異常音を聴取させる。

アルゴリズム:症例に対する対応を受験生が書き込む。

作問者への報酬: マルチメディアを導入するには作問者の能力とメディアの導入が必要であり、この場合には作問者への報酬が必要である。NBME では作問者への問題を返却して修正を強要できるのは報酬があるから？

2. その他

次回の開催は 2~3 月に中間報告を行う。

アメリカ医師国家試験委員会 (United States Medical Licensing Examination)

2015年11月4日
NBME Headquarters
Philadelphia, PA, USA

奈良信雄 (順天堂大学/東京医科歯科大学)
高木 康 (昭和大学)



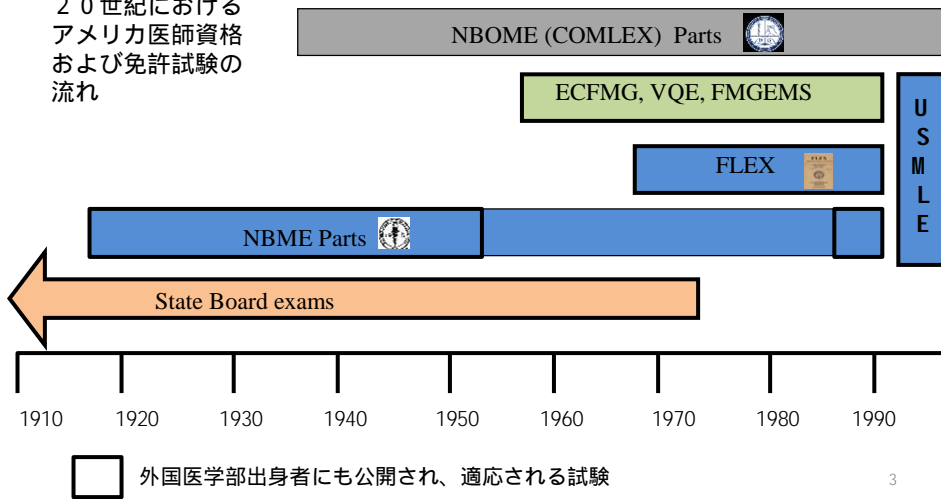
トピックス



- CBT導入の経緯
- CBTシステムの概略
- 試験方法
- 合否判定
- 管理・運営
- IT環境

USMLEの歴史

20世紀における
アメリカ医師資格
および免許試験の
流れ



USMLEのCBT実施経時的推移



1. USMLE 合意
2. USMLE 実施
3. NBME/FSMB 第1回CBT 問題配布 (SPEX)
4. USMLEにおけるCBT実施計画
5. USMLE CBT 開始
6. Step 2 Clinical Skills 追加



CBTの長所

- 紙と鉛筆による試験より機密保持に優れる。
 - ✓ 試験紙による試験ではコピーやカンニングがありうる
 - ✓ 試験用紙を配送中に盗難の虞あり
- CBT試験では、さまざまな機密保持の手段がある
- CBTでは、新しい試験法を導入可能である。
 - ✓ シミュレーション、ビデオ、音声を取り入れた試験など

5

CBTの短所

- USMLEのCBTは年間を通じて絶えず実施
 - ✓ 大量の試験問題をプールしないとイケない。
 - ✓ 試験実施に伴い、多数のスタッフが必要。

6

CBT試験システム

CBT試験システム

- 試験問題開発委員会が毎年新作問題を作成
- ブラッシュアップされた新作問題は非採点問題として出題される。
- 別の委員会により、統計学的に許容された問題がプールされる。
- プール問題に入れられた新作問題は毎年出題される。

一般的な問題作成法

作題者からNBMEへ問題提出

問題ブラッシュアップし、作題者へ返却

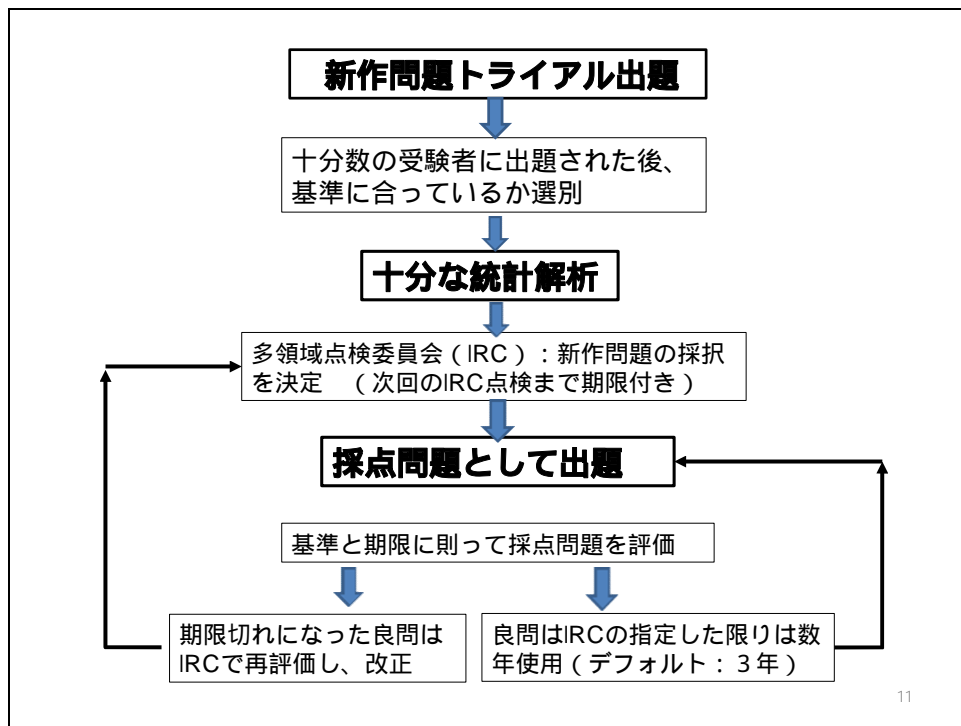
作題者による改変

委員会で全問題の確認と出題の決定

出題

新作問題の扱い

- 新作問題はトライアルとし、採点されない。
- 評価委員会で統計学に基づいて精査。
 - ✓ 良問として次回用にプール
 - ✓ 修正し、再度トライアル用
 - ✓ 削除
- 定期的に全問題をチェック (e.g., 陳腐な内容でないか; ガイドラインの変更に合致しているか)



CBT システム概要

- 試験規模と期間は受験者数に依る
 - ✓ 毎年各ステップの受験者は約30,000名
 - ✓ 試験はほぼ一年中試験会場で実施
- 各ステップでのテストフォームは数多く用意されている。

USMLE 運営規定

➤ 受験資格

- ✓ Steps 1 and 2
 - LCMEかAOAで認証を受けたアメリカ、カナダの医学部の学生または卒業生
 - ECFMGによる資格認定を受けたアメリカ、カナダ以外の医学部学生または卒業生
- ✓ Step 3
 - 医学部卒業生、整骨医学校卒業生、同等の資格者
 - Steps 1と2合格者
 - ECFMG資格を得たアメリカ、カナダ以外の卒業生

USMLE 運営方針

- **身体障害のある受験生への配慮**
 - ✓ 例：受験時間延長、筆記者の付き添い、音声記録
- **USMLE受験回数**
 - ✓ 受験は6回まで；医師免許機関の要請で例外あり
- **受験間隔**
 - ✓ 12か月以内の受験回数は3回まで
 - ✓ 4回目受験後の再受験は少なくとも6か月開ける

試験問題配信ソフト、 問題フォーマット

試験問題配信ソフト

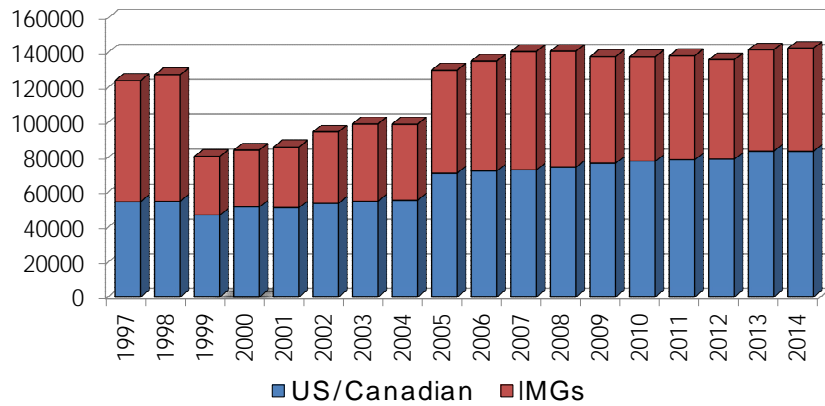
- 試験配信ソフトはNBMEが開発し、著作権もある。
- ✓ 2016年には拡大、反対色の新機能が導入される。
- 試験後のサーベイはSPSSソフトウェアで行う。

試験の関する フォーマットの提示

合否判定基準

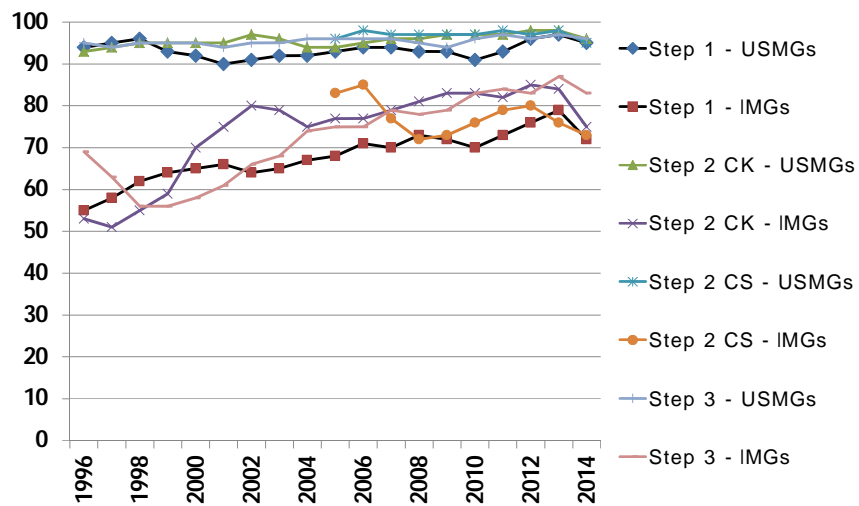
- 筆記試験とCBTは同じシステムである
- 基準判定を決定するグループの総括
 - ✓ 受験生のパフォーマンスの傾向
 - ✓ 調査を通してのステークホルダーの意見
 - ✓ Angoff法（修正アンゴフ法*1：試験の各設問ごとに、合否境界水準にある受験生のうち正解する者の割合を平均設定したものを、合議に基づき合否判定として設定する方法
 - ✓ 考えられる分類誤差の情報
- 合格最低レベルは3～4年毎に調査する

USMLE の受験者数の推移 (1997 ~ 2014)



19

初回受験者の合格点 (1996 ~ 2014)



結果報告

- 受験生に対して
 - ✓ 総得点、推奨される合否基準、分野別試験問題の長所と短所

21



UNITED STATES MEDICAL LICENSING EXAMINATION® STEP 2 CLINICAL KNOWLEDGE (CK) SCORE REPORT

This score report is provided for the use of the examinee.
Third party users of USMLE information are advised to rely solely on official USMLE transcripts.

Doe, Jane
USMLE ID: 0-000-000-0

Test Date: July 1, 2010

The USMLE is a single examination program consisting of three Steps designed to assess an examinee's understanding of and ability to apply concepts and principles that are important in health and disease and that constitute the basis of safe and effective patient care. Step 2 is designed to assess whether an examinee can apply medical knowledge, skills, and understanding of clinical science essential for the provision of patient care under supervision, including emphasis on health promotion and disease prevention. The inclusion of Step 2 in the USMLE sequence is intended to ensure that due attention is devoted to principles of clinical sciences and basic patient-centered skills that provide the foundation for the safe and competent practice of medicine. There are two components to Step 2: a Clinical Knowledge (CK) examination and a Clinical Skills (CS) examination. This report represents results for the Step 2 CK examination only. Results of the examination are reported to medical licensing authorities in the United States and its territories for use in granting an initial license to practice medicine. The two numeric scores shown below are equivalent; each state or territory may use either score in making licensing decisions. These scores represent your results for the administration of Step 2 CK on the test date shown above.

PASS

This result is based on the minimum passing score recommended by USMLE for Step 2 CK. Individual licensing authorities may accept the USMLE-recommended pass/fail result or may establish a different passing score for their own jurisdictions.

199

This score is determined by your overall performance on Step 2 CK. For recent administrations, the mean and standard deviation for first-time examinees from U.S. and Canadian medical schools are approximately 229 and 23, respectively, with most scores falling between 140 and 260. A score of 189 is set by USMLE to pass Step 2 CK. The standard error of measurement (SEM) for this scale is seven points.

22

INFORMATION PROVIDED FOR EXAMINEE USE ONLY

The Performance Profile below is provided solely for the benefit of the examinee. These profiles are developed as self-assessment tools for examinees only and will not be reported or verified to any third party.

USMLE STEP 2 CK PERFORMANCE PROFILE

	Lower Performance	Baseline Performance	Higher Performance
PHYSICIAN TASK PROFILE			
Preventive Medicine & Health Maintenance		XXXXXXXXXXXXXXXXXXXX	
Understanding Mechanisms of Disease	XXXXXXXXXXXX		
Diagnosis			XXXXXXXXXXXX
Principles of Management			XXXXXXXXXXXX
NORMAL CONDITIONS & DISEASE CATEGORY PROFILE			
Normal Growth & Development; Principles of Care		XXXXXXXXXXXXXXXXXXXX	
Immunologic Disorders		XXXXXXXXXXXXXXXXXXXX	
Diseases of Blood & Blood Forming Organs		XXXXXXXXXXXXXXXXXXXX	
Mental Disorders		XXXXXXXXXXXXXXXXXXXX	
Diseases of the Nervous System & Special Senses		XXXXXXXXXXXXXXXXXXXX	
Cardiovascular Disorders		XXXXXXXXXXXXXXXXXXXX	
Diseases of the Respiratory System		XXXXXXXXXXXXXXXXXXXX	
Nutritional & Digestive Disorders		XXXXXXXXXXXXXXXXXXXX	
Gynecologic Disorders		XXXXXXXXXXXXXXXXXXXX	
Renal, Urinary & Male Reproductive Systems	XXXXXXXXXXXX		
Disorders of Pregnancy, Childbirth & Puerperium		XXXXXXXXXXXXXXXXXXXX	
Musculoskeletal, Skin & Connective Tissue Diseases		XXXXXXXXXXXXXXXXXXXX	
Endocrine & Metabolic Disorders		XXXXXXXXXXXXXXXXXXXX	
DISCIPLINE PROFILE			
Medicine		XXXXXXXXXXXX	
Obstetrics & Gynecology		XXXXXXXXXXXXXXXXXXXX	
Pediatrics		XXXXXXXXXXXX	
Psychiatry		XXXXXXXXXXXXXXXXXXXX	
Surgery		XXXXXXXXXXXX	

The above Performance Profile is provided to aid in self-assessment. The shaded area defines a baseline level of performance for each content area; baseline performance is comparable to a HIGH FAIL/LOW PASS on the total test.

Performance bands indicate areas of relative strength and weakness. Some bands are wider than others. The width of a performance band reflects the precision of measurement; narrower bands indicate greater precision. The band width for a given content area is the same for all examinees. An asterisk indicates that your performance band extends beyond the displayed portion of the scale. Small differences in the location of bands should not be over interpreted. If two bands overlap, performance in the associated areas should be interpreted as similar. Because Step 2 CK is designed to be integrative, many items contribute to more than one content area. As a consequence, caution should be used when interpreting differences in performance across content areas.

This profile should not be compared to those from other Step 2 CK administrations.

Additional information concerning the topics covered in each content area can be found in the USMLE Step 2 CK Content Descriptions and Sample Test Materials.

結果報告

- **受験生に対して**
 - ✓ **総得点、推奨合格基準、試験の長所と短所**
- **もし受験生から以下のような要望がきたら**
 - ✓ **免許交付の決定における合格基準**
 - ✓ **医科大学に進級/卒業判定と悪いところの改善**
 - ✓ **卒後プログラムに対する入学基準の情報**

結果報告- 2

- 医科大学の全学生のデータ
 - ✓ 総数と合格率
 - ✓ 試験項目の長所と短所

25

NBME 2014年初受験でのパフォーマンス

Medical School:
School ID Number:

	PERFORMANCE ON FIRST ATTEMPT		PERFORMANCE ON MOST RECENT REPEAT ATTEMPT	
	Examinees from Your Medical School	Examinees from U.S. & Canadian Medical Schools	Examinees from Your Medical School	Examinees from U.S. & Canadian Medical Schools
Number Tested	102	22392	4	454
Number Passing	97	21474	4	352
Percent Passing	95	96	100	78
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Total Test	224 (20)	229 (20)	205 (7)	199 (14)

This report compares the performance of examinees from your medical school with the performance of examinees from U.S. and Canadian medical schools taking Step 1 of the United States Medical Licensing Examination (USMLE) for the first time during 2014. The performance of first-time examinees who failed Step 1 in 2014 but repeated the examination later in the year is also summarized above. Examinees who asked that their individual results not be provided to their school are excluded from the school data for this report, unless the number of examinees making this request was five or greater. The minimum passing score for the 2014 Step 1 administrations was 192.

Accompanying this table are two graphs: one that shows the distribution of scores (for your students and for the national first-taker group), and one that provides information regarding the performance of first-time examinees from your school for various disciplines and organ systems. Please note that the graphs are only provided if at least 20 first-time examinees from your school took Step 1 in 2014.

Step 1 test items are deliberately designed to be integrative; most items contribute to the calculation of subscores in more than one discipline. Consequently, caution should be used in attributing mean differences in student performance to individual courses at your school.

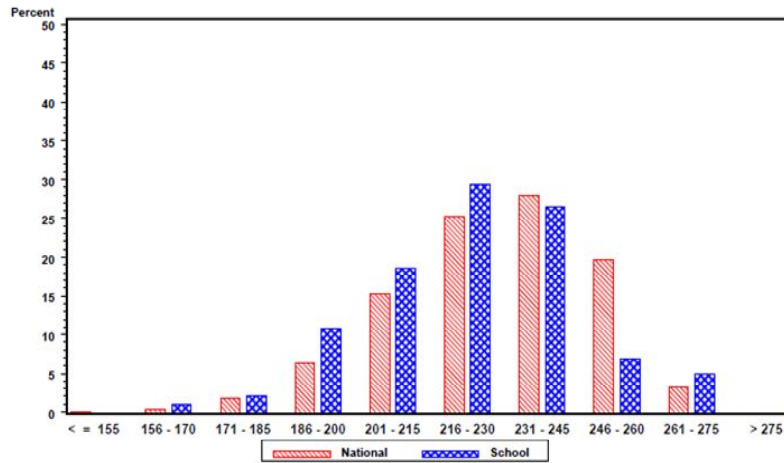
03102015

26

総得点の分布

2014年に初めてUSMLEを受験した学生

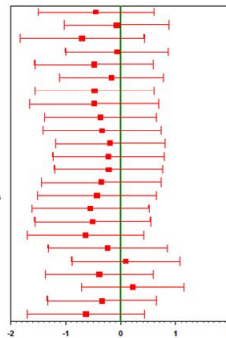
Medical School:



27

NBME2014年初受験でのパフォーマンス

- 1-Behavioral Sciences
- 1-Biochemistry
- 1-Biostatistics
- 1-Genetics
- 1-Gross Anatomy & Embryology
- 1-Histology & Cell Biology
- 1-Microbiology & Immunology
- 1-Nutrition
- 1-Pathology
- 1-Pharmacology
- 1-Physiology
- 2-General Principles of Foundational Science
- 2-Immune System
- 2-Blood & Lymphoreticular System
- 2-Behavioral Health & Nervous Systems/Special Senses
- 2-Musculoskeletal, Skin, & Subcutaneous Tissue
- 2-Cardiovascular System
- 2-Respiratory System
- 2-Gastrointestinal System
- 2-Renal/Urinary System
- 2-Reproductive System
- 2-Endocrine System
- 2-Multisystem Processes & Disorders
- 2-Biostatistics & Epidemiology/Population Health

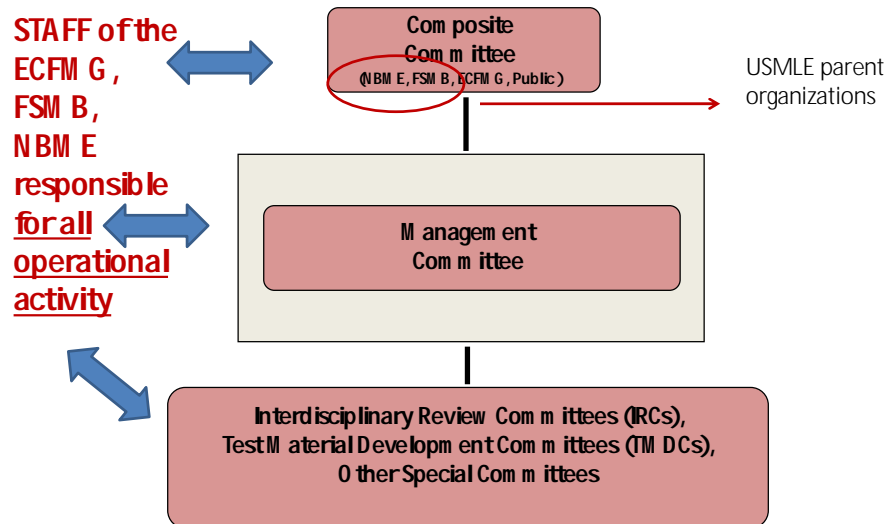


The above graph provides information regarding the score distribution of first takers from your medical school relative to the distribution for all U.S./Canadian first takers in each discipline and organ system. All scores are scaled in standard score units based on the performance of U.S./Canadian first takers: the mean and standard deviation (SD) for this group are 0 and 1, respectively, for each discipline and organ system. To facilitate interpretation, the reliability of each score category has been used in adjusting the standard scores. This adjustment helps to make the differences in standard scores a better reflection of true differences in student performance. The mean performance of U.S./Canadian first takers is represented by the vertical solid green line at 0.0. Roughly 68% of U.S./Canadian first takers scored within one SD of the mean, between -1.0 and 1.0. The distribution of performance for first takers from your school is represented by the red boxes and horizontal lines. The red box depicts the mean performance of first takers from your school. The distance from the red box to one end of the red line indicates one SD for your school. The interval spanned by each red line represents your school mean plus/minus one SD, approximately 68% of your students scored in this interval.

By comparing the locations of the red boxes, you can determine the disciplines and organ systems in which the performance of your students was relatively strong or weak. Because many of the scores are based on a relatively small number of items, differences smaller than a few tenths of an SD are not likely to be meaningful. In addition, because Step 1 test items are deliberately designed to be integrative with many items contributing to the calculation of scores in more than one discipline, caution should be used in attributing mean differences in student performance to individual courses at your school. Content areas are coded based on two dimensions as follows: 1 - Discipline, and 2 - System.

28

USMLEの管理システム



29

受験料 (2015) (米国ドル)

試験	米国学生	米国/カナダ以外
Step 1	590	850
Step 2 CK	590	850
Step 2 CS	1,250	1,480
Step 3	815	815

****米国やカナダ以外の地域で受験する場合には地域により別途料金がかかる**

30

USMLEに関する情報

- USMLE Website (www.usmle.org)
 - ✓ 試験に関する情報、試験サンプル、運営規則、試験に関するデータ、解釈のガイド、引用



第 4 回研究会議

日時:平成 28 年 2 月 18 日(木) 10:00～12:00

場所:公益社団法人 医療系大学間共用試験実施評価機構 6 階会議室

参加者:石田達樹、大西弘高、鈴木利哉、奈良信雄、高木 康

アドバイザー:鈴木貴士専門官

議題

1. 平成 27 年活動の振り返り

1) 第 1 回研究会議(平成 27 年 6 月 23 日)

研究会議の目的:医師国家試験の在り方に関する研究

マルチメディア素材を用いた CBT の可能性を探る

- ・ あくまでも全国同時実施の形式は堅持する方向での検討
- ・ プール化は現状では想定しなくてよい
- ・ メリット、デメリットを複数の視点議論する
- ・ 諸外国での取り組みについて現地調査を行い参考にする

2) 第 2 回研究会議(平成 27 年 8 月 13 日)

マルチメディアを用いた CBT の可能性について識者からの講演

岡崎仁昭先生(自治医科大学)

- ・ 動画素材を CBT に導入(自治医科大学、内科学会)
- ・ 神経所見・心エコーなどは動画を用いることで深い知識を問うことが可能である。

青木茂樹先生(順天堂大学)

- ・ 放射線映像素材の導入
- ・ キーフィルム以外も含め、異常所見画像を受験者が見つけるようなパターンも考慮する。

台湾の国家試験

- ・ 台湾の医師国家試験パンフレット
- ・ 奈良委員、高木委員からの報告

3) 第 3 回研究会議(平成 27 年 12 月 11 日)

米国の国家試験

- ・ 奈良委員、高木委員からの報告
- ・ NBME、ECFMG 提供の PPT

今後

- ・ 次年度も同様な課題で申請して補助を受け、研究を継続する。

- ・ 次年度はプロトタイプの CBT 問題を試行する可能性を探る。

2. 報告書作成

1) 報告書内容

班会議の議事録
台湾と米国の視察報告書
岡崎先生と青木先生提供のスライド
インドネシアの試験事情(大西委員)

3. 発展的 CBT について

1) マルチメディア対応 CBT の可能性

現状の冊子式の国家試験では情報量が限られているが、マルチメディア対応 CBT では設問において受験生が読み取る情報量を多くすることが可能である。

CBT では設問を 1 画面上に設定するがタブによる他画面参照も考慮する。

CBT では単純 5 選択肢択一形式と多選択肢択一形式では情報量が異なる。

設問ごとの情報量が異なる場合には解答時間を調整する必要がある。

日本語診療能力試験など少人数での試験に試行的に導入することも可能か。

2) マルチメディアを利用する CBT の注意点

画像の場合は解像度を適切に設定する必要がある・試験に使用する端末の解像度が 4K かフルHDかによっても対応が異なるが、受験生が読み取れるかが重要である。心エコー図など素材ごとに適切な解像度を設定する必要がある。動画については使用する規格を設定する必要がある。

国家試験の画像では全受験生が同じ解像度である必要がある。

全ての問題にマルチメディアを利用する必要はない

共用試験のように試験ブロック管理を採用し、特定のブロックだけをマルチメディア利用問題とすることも可能か。

試行導入においては 10 題程度を CBT 形式で動画・音声を利用した設問とする。

解像度は共用試験で採用する解像度あるいは日常診療で使用している程度で良いか。

- ・ いわゆるプロ(専門医)が使うようなものではなく、典型例でよい
- ・ 情報量はさほど高いものである必要はない(画質は 720 × 480 など)

従来の問題よりも作問等に手間がかかる可能性が高い

作問者への報酬等も考慮すべきか検討が必要である。NBME では報酬制をとっている。

デメリット: 一度使用した問題は再度使用できない可能性が高い。

- ・ プール問題制はとれず、使い捨てしかできない。
- 最終的にはマルチメディアを導入した CBT をパイロット的に試行する。
試験を実施管理するサーバをクラウドにするかローカルにするか
- ・ クラウドサーバにすると大学の負担は減る。
 - ・ 直接クラウドサーバにアクセスするならネットの通信速度や安定性などに依存する。
 - ・ 台湾の医師国家試験においては中央サーバ・ローカルサーバ・試験クライアントの三層構造になっている。
 - ・ クラウドサーバを利用した場合、クラウドサーバと試験実施大学の間のネットワークが切断された場合に試験が中断する可能性がある。
- 常設の PC か、個人の PC か
- ・ BYODにどのように対応していくか見当が必要である。試験センターの活用も検討する必要がある。
 - ・ マルチメディア素材を扱うことを前提にした場合、一定のスペックを要求する必要がある。

3) 動画や音声を CBT に利用する場合

音声(心音、呼吸音など)と動画(心エコー、患者身体診察など)を CBT に利用

音声はそれほど容量が大きくないので可能か

画像は種類によるが容量の差異が大きすぎて利用できない場合もある。

配信のパターン(試験前、試験間、試験問題毎)について検討が必要。

端末機器ごとに選択できるような方式を採用すべき

画像の拡大機能(台湾では採用)

- ・ ズームに対応した画像を用意しないと判読しにくくなる。
- ・ スケールを受験生が自由に決定できるのか、あるいは×2、×4 倍のように固定倍率を規定しておくか

4) 問題解決型シミュレーション形式の問題

電子カルテに掲示されているような情報をクリックして問題解決を行う形式にする

既往歴、画像は電子カルテと同様にあるが、すべてを見ると時間が足りない。実際の診療では有効的である。

研修医1年目が診療上の情報過多で困っているのであれば、そういう問題を出題することで学習を促すことができる

時間軸を含めた課題、時間制限の設定も可能

受験生が選択した画像が入手できるシステムのするか

解答は記入式、選択肢(MCQ や EMI)のいずれか

キーワードを散りばめる形ではない問題

ハイパーリンクやクリックابلマップの機能も技術としては使用できる。

5) 試行 CBT 問題

現行の医師国家試験形式に適応するには専門的な知識を有した技術者に相談する必要がある。

日本語診療能力試験に適応する？ ブロックとしてマルチメディアを利用した CBT を作成した試用する。この場合には PC もそれほど必要としない。

日本語診療能力試験には CBT を導入することは考えているが、マルチメディアを利用した試験を試用することはできない(鈴木専門官)

試用は可能であるが、本格導入は難しいか。

学生に試用して、通常の CBT、OSCE との相関を検討する必要があるが、問題数は限られるので、評価は難しいか

専門医・専攻医試験でも利用できるか

海外でもマルチメディア対応は行っていないのは導入が難しい

4. 次年度厚生労働省科学研究費の件

今年度の研究を継続する内容の研究費の申請を行う予定である。

5. その他