

厚生労働行政推進調査事業費補助金（地域医療基盤開発推進研究事業）
総括研究報告書

基本的臨床能力評価試験の質向上についての研究
(21IA2004)

研究代表者 西崎 祐史
順天堂大学 医学部医学教育研究室 先任准教授

〔研究要旨〕

本研究班では、臨床研修医の客観的臨床能力評価のために開発された「基本的臨床能力評価試験（GM-ITE: General Medicine In-Training Examination）」の質向上を主な目的とし、「GM-ITE 問題作成プロセスのブラッシュアップ」、「CBT（Computer-Based-Testing）導入による試験問題管理の効率化」、「実践経験の評価に即した問題作成」、「GM-ITE バリデーション（GM-ITE の国際的妥当性等の検証）」、「研修医労働時間及び研修環境と基本的臨床能力評価試験等との関連性の検討」、「新旧臨床研修プログラムの比較」の6つのプロジェクトを同時進行している。

「GM-ITE 問題作成プロセスのブラッシュアップ」については、2022 年度問題作成に向けて、11 回のウェブ会議を行い、問題作成委員会を中心にのべ 60 名が会議に参加した。2021 年度の試験結果データを基に、本年度試験の出題分野や難易度等について、活発な議論がなされ、問題作成プロセスのブラッシュアップが図られた。「CBT 導入による試験問題管理の効率化」については、2021 年度に引き続き Moodle を使用した CBT を実施し大きな問題はなかった。試験の更なる質向上を目指し、厚労科研「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（河北班）」と会議（2022 年 8 月 3 日）を行い、主に CBT 運用面について意見交換を行った。「実践経験の評価に即した問題作成」については、56 名の研修医がパイロット的に受験した患者再現 VTR 問題の試験結果を解析した。その結果、本問題は「総論（医療面接・プロフェッショナリズム）」、「症候学・臨床推論」、「身体診察法・臨床手技」、「疾病各論」の4つのドメインで高い識別指数を示した。「GM-ITE バリデーション（GM-ITE の国際的妥当性等の検証）」については、タイのマヒドン大学の研究グループとの国際共同研究が進捗した。具体的には、タイにおける倫理委員会での承認が得られ、タイの医学生と研修医を対象とした国際比較研究実施の準備が整った。「研修医労働時間及び研修環境と基本的臨床能力評価試験等との関連性の検討」については、研修医のメンタルヘルスという観点では、うつ、バーンアウト、高ストレスの有病率は週平均労働時間が 90 時間以上で有意に増加することが分かった。

新型コロナウイルス感染症診療に関しては、2021 年 2 月に実施した研修環境調査アンケート（n=5,976）の結果から、47%（2,807 名）の研修医が新型コロナウイルス感染症診療を経験していないことが分かった。また、1 年後（2022 年 2 月）に実施したアンケート（n=3,869）の結果からは、30.9%（1,195 名）の研修医が新型コロナウイルス感染症診療未経験であることが分かった。さらに、過剰な新型コロナウイルス感染症診療経験は研修医のバーンアウトと有意な関連性を示し、個人防護具（Personal Protective Equipment; PPE）の十分な支給が、研修医バーンアウト罹患割合と、負の関連性を示した。「新旧臨床研修プログラムの比較」については、解析可能なデータ規模の拡張を目的とした研究計画変更の申請を行った。倫理委員会における審査により、2022 年 9 月に研究計画変更が承認された。

GM-ITE は現在、研修医の約半数が受験しているが、本研究班の活動を通じて、全研修医が受験するようになれば、日本全国の臨床研修施設における研修プログラムの評価が可能となり、臨床研修医教育の標準化および質向上が実現し、延いては日本の医療の質向上に繋がる。また、本研究班は GM-ITE の結果及び研修環境調査アンケートの結果を活用し、至適な研修教育環境に関するエビデンスを発信し続けている。本研究班から発信されるエビデンスが、研修医教育に関する医療政策に資する資料となることが期待される。

〔研究分担者〕

徳田安春 群星沖縄臨床研修センター 群星沖縄臨床研修センター センター長
志水太郎 獨協医科大学 総合診療医学 主任教授
山本祐 自治医科大学 地域医療学センター総合診療部門 講師
鋪野紀好 千葉大学 地域医療教育学 特任准教授
福井翔 杏林大学 総合医療学 助教
小林裕幸 筑波大学 医学医療系 教授
長崎一哉 筑波大学付属病院 水戸地域医療教育センターJA 茨城県厚生連総合病院水戸協同病院
総合診療科 医師
西口翔 横浜市立大学 公衆衛生学教室 客員准教授
片山皓太 聖マリアンナ医科大学 総合診療内科 助教

〔背景・目的〕

2004 年 4 月より卒後臨床研修が義務化され、スーパーローテーションが実施されるようになったが、研修プログラムの運営や実施体制は各医療機関の裁量に委ねられている所が大きい。さらには、臨床研修における客観的アウトカム指標は確立されておらず、教育内容は医療機関によって様々である。その結果、研修医のスキルにも大きな差が生まれているのが現状である。このような本邦の現状において、研修医教育の標準化および質向上は喫緊の課題である。それらの課題解決を目指し、2005 年 9 月に NPO 法人日本医療教育プログラム推進機構（JAMEP: Japan Institute for Advancement of Medical Education Program）が設立された。

研修医教育の標準化および質向上を達成するためには、臨床研修を客観的に評価するための指標が必要である。米国においては、米国医師資格試験（USMLE: United States Medical Licensing Examination）により臨床研修の客観的評価が可能であるが、本邦においては確立された評価指標は存在しない。

そこで、JAMEP は基本的臨床能力の客観的な評価指標として「基本的臨床能力評価試験（GM-ITE）」を開発した。GM-ITE は、臨床研修医を対象とした「In-Training Exam」であり、2011 年度（第 1 回）より導入され、2022 年度（第 12 回）には 662 病院、約 9,011 名の研修医が参加するまでの規模に拡大した。問題は「総論（医療面接・プロフェッショナリズム）」、「症候学・臨床推論」、「身体診察法・臨床手技」、「疾病各論」の 4 分野で構成されており、幅広い疾患領域（内科・外科・小児科・産婦人科・救急科・精神科等）が網羅されている。

GM-ITE は行政的な観点で、重要な位置を占める。具体的には、令和元年 7 月 3 日に開催された、第 2 回医道審議会医師分科会医師臨床研修部会において、「臨床研修プログラム評価・客観的な研修医の能力評価のために JAMEP の基本的臨床能力評価試験等の客観的な能力試験の活用を推奨することとしてはどうか。」という提案がなされた。臨床研修医の評価は、指導医による到達目標に基づく評価や 360 度評価に加え、客観的な能力試験を用いた多角的な方法が望まれている。

本研究班は、GM-ITE の質向上に必要な要素を抽出し、課題解決を通じて試験の質向上を目指す。

〔方法〕

本研究班では、GM-ITE の質向上を主な目的とし、「GM-ITE 問題作成プロセスのブラッシュアップ」、「CBT 導入による試験問題管理の効率化」、「実践経験の評価に即した問題作成」、「GM-ITE バリデーション（GM-ITE の国際的妥当性等の検証）」、「研修医労働時間及び研修環境と基本的臨床能力評価試験等との関連性の検討」、「新旧臨床研修プログラムの比較」の 6 つのプロジェクトを同時進行している。

① GM-ITE 問題作成プロセスのブラッシュアップ

GM-ITE の試験問題は、総合医と専門医から構成される問題作成委員会により作成される。作成された問題は、毎年、査読委員会での査読や、シニアレジデントによるパイロット試験、テスト分析により、客観性が担保される。本プロジェクトでは、試験問題分野、問題作成委員および査読委員等の人数の適正化および選抜方法の見直し等の問題作成プロセスのブラッシュアップ方法を検討する。また、プール問題の活用方法等の検討を行う。

② CBT 導入による試験問題管理の効率化

受験者数の増加に伴い、受験生の利便性向上および、試験問題管理の効率化を目的とし、2021 年度試験から CBT を導入した。CBT 導入により、従来型筆記試験では不可能であった設問が可能となるため、試験問題の質向上が可能となる。本プロジェクトでは、CBT 試験運営等の更なる効率化を目指す。

③ 実践経験の評価に即した問題作成

GM-ITE は、臨床研修の実践で培われるスキルが評価できるように、医療現場でのマネジメントを問う問題が多く含まれる。2018 年度試験からは、身体所見を中心とした動画問題を導入した。本プロジェクトでは、動画問題の増加に加え、患者再現 VTR を用いた問題等の開発を進め、より実践経験の評価に即した問題の作成を目指す。

④ GM-ITE バリデーション（GM-ITE の国際的妥当性等の検証）

GM-ITE では、毎年、全体の 20%(12/60 問)は英語問題として出題している。過去に出題した英語問題の中で、正答率や識別性等を評価したテスト解析で良問と判断された問題を抽出し、アジア諸国の研修医等を対象に試験を実施する。なお、新型コロナウイルス感染症の収束が見通せない状況下では、海外の研究者との連携は、対面に限らず、ウェブやメール会議を通じて行う。海外の研修医等の試験結果データから識別指数等を計算し、GM-ITE の国際的妥当性を検証する。また、諸外国の研修医や医学生と日本の研修医のスコアを比較することで、日本の研修医の学習到達レベルが国際的にどの位置にいるのかを問題領域別に把握する。

⑤ 研修医労働時間及び研修環境と基本的臨床能力評価試験等との関連性の検討

研修医の至適労働時間を決定していく上で、労働時間と基本的臨床能力の関連性の検討は重要である。本研究班は、日本最大規模のデータ（約 6,000 名）を用いた解析を実施し、基本的臨床能力開発の観点からは、臨床研修医の至適労働時間は、週平均 60-65 時間（年平均時間外労働時間としては 960-1,200 時間）であり、研修効果を最大化する観点からは C-1 水準の時間外労働時間の上限年 1,860 時間は、今後は減少させることが適当であるというエビデンスを示し、その結果を、厚生労働省第 10 回医師の働き方改革の推進に関する検討会（2020 年 11 月 18 日）で報告した。本プロジェクトでは、研修医を対象に実施している研修環境調査アンケートのデータを活用し、様々な研修環境等と GM-ITE スコアとの関連等を検討したり、研修医のメンタルヘルスや医療安全に関連する因子（研修医の労働時間を含む）等の評価を行う。これらの評価は、安全で質の高い指導環境の構築を目指す上で重要な基礎資料と成り得る。

⑥ 新旧臨床研修プログラムの比較

2020 年度に臨床研修の内容が、到達目標を中心に大幅に見直された。改訂後の新研修プログラムは、その教育効果について、評価・検証される必要がある。GM-ITE 結果を活用し、改訂後の新研修プログラムの評価・検証が実施されれば、研修プログラムの更なる改善に繋がる基礎資料が得られ、延いては、国民の健康増進に繋がる。

<倫理的配慮>

本研究で実施される研究は、倫理審査が必要な研究テーマについては、倫理委員会の承認を得て実施されている。また、現在計画中の研究においても、倫理委員会の承認を得た後に実施する予定としている。また、研究の目的及び説明文を GM-ITE に参加した研修医個人に提示し、オンライン上で試験結果及びアンケート結果の研究利用について同意を得る。また、JAMEP ホームページ上に、研究参加の拒否（オプトアウト）を受け入れることを明記しており、希望者は手続きが可能である。研究実施後も、個人情報の保護等や研究データ管理については、厳重に管理する。また、研究の目的以外に、研究で得られた研究対象者の試料等を使用しない。

[結果]

① GM-ITE 問題作成プロセスのブラッシュアップ

2022 年度問題作成に向けて、11 回のウェブ会議を行い、問題作成委員を中心にのべ 60 名が会議に参加した。昨年度の試験結果データを基に、本年度試験の出題分野や難易度等について、活発な議論がなされ、問題作成プロセスのブラッシュアップが図られた。その結果、2022 年度 GM-ITE には、全国 662 施設、9,011 人の研修医が参加し、2021 年度の受験者数を上回った。

② CBT 導入による試験問題管理の効率化

2022 年度も Moodle を使用し、CBT を実施したが、大きな問題はなかった。試験の更なる質向上を目指し、2021 年度に引き続き、厚労科研「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（河北班）」と会議（2022 年 8 月 3 日）を行い、主に CBT 運用面について意見交換を行った。今後も連携を継続する予定である。

③ 実践経験の評価に即した問題作成

肺血栓塞栓症の症例を基にした患者再現 VTR 問題を作成し、非言語的コミュニケーション能力を含む評価を試みた。56 名の研修医が患者再現 VTR 問題に解答し、本試験問題の有用性が示された。具体的には、患者再現 VTR 問題における、診断名を問う設問については、GM-ITE 全体で 0.47 と高い識別指数を示した。さらに「総論（医療面接・プロフェッショナリズム）」、「症候学・臨床推論」、「身体診察法・臨床手技」、「疾病各論」の全てのドメインで、識別指数 0.35 以上と比較的高い数値を示した。

また、SBAR（状況・背景・評価・提案）についての自由記載解答において、医師による採点（2 名で採点）と機械による採点（形態素解析）を行ったところ、両者の差は 7.2% であり、両群での総得点に有意差はなかった（ $p=0.683$ ）。また、点数の相違が出た理由には、人的採点によるエラーや解答の誤字が挙げられた。

④ GM-ITE バリデーション（GM-ITE の国際的妥当性等の検証）

タイのマヒドン大学の研究グループとの国際共同研究が進んでいる。2022 年度には、タイにおける倫理委員会での承認が得られた。タイの医学生と研修医を対象に、過去に出題された GM-ITE 英語問題の中で、識別指数の高い 50 問の試験実施が予定された。2023 年度は、結果の解析を通じ、日本の研修医の学習到達レベルが国際的にどの位置にいるのかを問題領域別に把握する予定である。

⑤ 研修医労働時間と基本的臨床能力評価試験の関連性の検討

2021 年度、GM-ITE スコアと労働時間の関連性を検討した論文が、Medical Teacher 誌に掲載されたが、それに続き、2022 年度は研修医のメンタルヘルスと労働時間との関連性を評価したエビデンスが Scientific Reports 誌に掲載された。具体的には、「研修医のメンタルヘルスという観点では、うつ、バーンアウト、高ストレスの有病率は週平均労働時間が 90 時間以上で有意に増加する」という内容を報告した。（Sci Rep. 2022. PMID: 35739229）

また、新型コロナウイルス感染症診療に関しては、2021 年 2 月に実施した研修環境調査アンケート（ $n=5,976$ ）の結果から、47%（2,807 名）の研修医が新型コロナウイルス感染症診療を経験していないことが分かった。また、1 年後（2022 年 2 月）に実施したアンケート（ $n=3,869$ ）の結果からは、30.9%（1,195 名）の研修医が新型コロナウイルス感染症診療未経験であることが分かった。さらに、過剰な新型コロナウイルス感染症診療経験は研修医のバーンアウトと有意な関連性を示し、PPE の十分な支給が、研修医バーンアウト罹患割合と、負の関連性を示した。（BMJ Open. 2023. PMID: 36639218）

⑥ 新旧臨床研修プログラムの比較

GM-ITE データ（過去問題データを含む）を用いて、新研修プログラム下での研修医と旧プログラム下での研修医の GM-ITE 成績を比較し、新旧臨床研修プログラムの比較検討を進められた。解析可能なデータ規模の拡張を目的とした研究計画変更の倫理審査が 2022 年 9 月に承認された。

[考察]

2022 年度の GM-ITE には、全国 662 施設、9,011 名の研修医が参加し、受験者数が、本邦の臨床研修医全体の約半数に到達するまでに拡大してきている。GM-ITE の問題作成に向けた問題作成委員会等の体制強化、CBT 導入による効率化、動画問題や患者再現 VTR 問題を中心とした実践経験の評価に即した試験問題の充実化、GM-ITE の妥当性の検証等に基づいた様々な取組みが受験者数の増加に結びついたと考える。

GM-ITE は、国全体での客観的研修医評価方法のモデルとしても重要な位置を占める。2019 年 7 月 3 日に開催された、第 2 回医道審議会医師分科会医師臨床研修部会において、「臨床研修プログラム評価・客観的な研修医の能力評価のために JAMPEP の基本的臨床能力評価試験等の客観的能力試験の活用を推奨することとしてはどうか。」という提案がなされた。研修医の評価は、指導医による到達目標に基づく評価

や 360 度評価に加え、GM-ITE 等の客観的能力試験を用いた多角的な方法が望まれている。本研究班の活動により、客観的な臨床研修プログラムの評価及びフィードバック法が確立されれば、本邦の臨床研修医教育の標準化及び質向上が実現する。

本研究班から発信されるエビデンスは、医師の働き方改革においても重要な資料と成り得る。2024 年 4 月より、医師の時間外労働を原則的に年 960 時間かつ月 100 時間までとする基準（A 水準）が設けられることとなった。ただし、地域医療や救急体制の維持に関わる医療機関（B 水準）や、研修医（C-1 水準）、高度技能の育成が公益上必要な分野について特定の医療機関で診療に従事する場合（C-2 水準）に関しては、年 1,860 時間かつ月 100 時間までの時間外労働が許容された。しかしながら、C-1 水準を許容すべきかどうかについては議論の余地がある。そこで、本研究班は、日本最大規模のデータ（約 6,000 名）を用いた解析を実施し、基本的臨床能力開発の観点からは、臨床研修医の至適労働時間は、週平均 60-65 時間（年平均時間外労働時間としては 960-1,200 時間）であり、研修効果を最大化する観点からは C-1 水準の時間外労働時間の上限年 1,860 時間は今後減少させることが適当であるというエビデンスを示し、その結果を、厚生労働省第 10 回医師の働き方改革の推進に関する検討会（2020 年 11 月 18 日）で報告した。

さらに、2022 年度には、うつ、バーンアウト、高ストレスの有病率は週平均労働時間が 90 時間以上で有意に増加するという内容を報告した。本結果は、研修医の基本的臨床能力の開発とメンタルヘルスの保護のバランスを考える上で重要な資料と成り得る。

新型コロナウイルス感染症の拡大に伴い、研修医の研修機会が制限される場面を目にするようになった。その理由は、院内の感染管理と、研修医のバーンアウト予防が目的であると考えられる。研修医の新型コロナウイルス感染症診療機会をコントロールすることで、病院における感染管理の強化や研修医のメンタルヘルスは守られる。その一方で、新型コロナウイルス感染症診療を通じた感染症診療経験の制限は、研修医の基本的臨床能力の開発の妨げになる可能性がある。このような、状況で実施した、研修医の新型コロナウイルス感染症診療実態調査から、約半数の研修医が新型コロナウイルス感染症診療未経験であることが判明した。感染管理やメンタルヘルスの観点から、保守的な選択がなされているのだろう。しかしながら、研修医の感染症診療における臨床能力開発という観点では、今の状況は適切であるとは言えない。指導医によるサポート、適切な PPE の支給を徹底した上で、研修医の新型コロナウイルス感染症診療機会の調整を検討するべきである。

これらのエビデンスは、研修医の働き方改革や新興感染症流行時の研修医教育の在り方を検討する上で、貴重な資料となる。今後も、本研究班は、GM-ITE の結果及び、GM-ITE 受験後に研修医を対象に実施している研修環境調査アンケートの結果を活用し、至適な研修教育環境に関するエビデンスを発信し続けていく。本研究班から発信される情報が、研修医教育に関する医療政策に資する資料となることが期待される。

今後も引き続き、「GM-ITE 問題作成プロセスのブラッシュアップ」、「CBT 導入による試験問題管理の効率化」、「実践経験の評価に即した問題作成」、「GM-ITE バリデーション（GM-ITE の国際的妥当性等の検証）」、「研修医労働時間及び研修環境と基本的臨床能力評価試験等との関連性の検討」、「新旧臨床研修プログラムの比較」の 6 つのプロジェクトをベースに、研究を進めていく予定である。

また、今後は、GM-ITE 受験後に研修医を対象に実施している研修環境調査アンケートの結果を活用し、まだ評価できていない研修環境要因と GM-ITE スコアとの関連を検討したり、研修医のメンタルヘルスや医療安全等に関連する要因を検討していく。これらの評価は、安全で質の高い研修指導環境の構築を目指す上で重要な基礎資料と成り得る。

[結論]

2022 年度 GM-ITE には、全国 662 施設、9,011 名の研修医が参加し、前年度の受験者数を上回った。GM-ITE の問題作成に向けた問題作成委員会等の体制強化、CBT 導入による効率化、動画問題や患者再現 VTR 問題を中心とした実践経験の評価に即した試験問題の充実化、GM-ITE の妥当性の検証等に基づいた様々な取り組みが、受験者数の増加に繋がった。

また、研修医のうつ、バーンアウト、高ストレスの有病率は週平均労働時間が 90 時間以上で有意に増加する可能性が示された。研修指導責任者は、研修医の基本的臨床能力の開発とメンタルヘルスの保護の両方のバランスを取りながら研修指導に取り組む必要がある。

[健康危険情報]

特記すべき事項なし。

[研究発表]

1. 論文発表

・Nagasaki K, Nishizaki Y, Shinozaki T, Shimizu T, Yamamoto Y, Shikino K, Fukui S, Nishiguchi S, Kurihara M, Kataoka K, Tokuda Y, Kobayashi H. Association between mental health and duty hours of postgraduate residents in Japan: a nationwide cross-sectional study. Sci Rep. 2022; 12(1): 10626. PMID: 35739229.

・Nagasaki K, Nishizaki Y, Hachisuka C, Shinozaki T, Shimizu T, Yamamoto Y, Shikino K, Fukui S, Nishiguchi S, Katayama K, Kurihara M, Kobayashi H, Tokuda Y. Impact of duty hours on competency-related knowledge acquisition among community hospital residents. J Gen Fam Med. 2022; 24(2): 87-93. PMID: 36909787.

・Nishizaki Y, Nagasaki K, Shikino K, Kurihara M, Shinozaki T, Kataoka K, Shimizu T, Yamamoto Y, Fukui S, Nishiguchi S, Katayama K, Kobayashi H, Tokuda Y. Relationship between COVID-19 care and burnout among postgraduate clinical residents in Japan: a nationwide cross-sectional study. BMJ Open. 2023; 13(1): e066348. PMID: 36639218.

2. 学会発表

・西崎祐史（座長）、山本祐（演者）、鋪野紀好（演者）、志水太郎（演者）、長崎一哉（演者）、医学教育（Medical Education）Vol.53, suppl. 2022（第54回日本医学教育学会大会予稿集）、シンポジウム4 本邦の初期臨床研修医教育における基本的臨床能力評価試験（GM-ITE: General Medicine In-Training Examination）の活用

・長崎一哉、医学教育（Medical Education）Vol.53, suppl. 2022（第54回日本医学教育学会大会予稿集）、優秀発表賞（Young Investigator Award）YIAS-3 臨床研修医の労働時間は自習時間にどのような影響を与えるのか？：全国横断研究（YIA 演題として表彰）

・松本貴文、長崎一哉、西崎祐史、篠崎智大、志水太郎、山本祐、鋪野紀好、福井翔、西口翔、栗原健、小林裕幸、徳田安春、医学教育（Medical Education）Vol.53, suppl. 2022（第54回日本医学教育学会大会予稿集）、口演 18 0-18-2 臨床研修医の労働時間は満足度にどのような影響を与えるのか？：全国横断研究

3. シンポジウム

- ・JAMEP 基本的臨床能力評価試験シンポジウム 2022, 2022 年 9 月 4 日, ウェブ開催
プレイベント「座談会」（司会：徳田安春、西崎祐史）
第1部「2020 年度「基本的臨床能力評価試験」総括」（演者：西崎祐史）
第2部「成績上位医療機関における臨床研修の取り組み」（演者：西崎祐史）
第3部「パネルディスカッション」（司会：西崎祐史）

[知的財産権の出願・登録状況]

特記すべき事項なし。