

医師国家試験出題基準に関する研究
～医学総論および必修の基本的事項の検証～

研究分担者 小松 弘幸(宮崎大学医学部 医療人育成推進センター 教授)
小西 靖彦(順天堂大学医学教育研究室 特任教授)
錦織 宏(名古屋大学医学系研究科 教授)
伴 信太郎(愛知医科大学医学教育センター 特命教育教授)
高村 昭輝(富山大学学術研究部医学系 医学教育学 教授)

研究要旨

医師国家試験改善検討部会報告書に示された「医師国家試験の出題基準を可能な限り臨床実習における経験を評価する内容に絞るよう見直しを行うことが妥当」の提言に基づき、卒前・卒後医学教育部門に属する教員(医師)18 名を検証パネラーとして、第 118 回医師国家試験の医学総論および必修の各問題について4段階のレベル評価(I:臨床実習前の学修で回答可能、II:臨床実習中の成果を問う、III:臨床研修で修得すべきもの、IV:専門研修で修得すべきもの)を行った。その結果、レベルIIと判断された設問は、医学総論の一般問題で45%、臨床実地問題で73%であり、必修の一般問題で41%、臨床実地問題で69%であった。一方、レベルIと判断された設問は、それぞれ医学総論の一般問題で47%、臨床実地問題で16%、必修の一般問題で55%、臨床実地問題で24%であった。パネラーの半数以上がレベルIと判定した問題の割合は、医学総論および必修問題の双方で一般問題が臨床問題より有意に高かった(医学総論:53% vs. 6%, $p < 0.01$ 、必修:52% vs. 2%, $p < 0.01$)。

医師国家試験における「医学総論」は平成5年(第87回試験)、「必修の基本的事項」は平成9年(第91回試験)に導入され、約30年の経過の中で平成13年の医学教育モデル・コア・カリキュラム策定、平成16年の臨床研修制度開始、平成18年の共用試験全国実施など、医学教育を取り巻く状況は大きく変化してきた。医師国家試験における現状の3区分(必修、医学総論、医学各論)が今後も独立して必要かどうかを含めた、医師国家試験全体のあるべき像を描く時期である。その際には、医師国家試験と医学教育モデル・コア・カリキュラム、臨床実習前共用試験との整合をとり、医学生に対する一貫した学修者評価のブループリントを策定することが望まれる。

A. 研究目的

1. 「医学総論」および「必修の基本的事項」の導入とその後の経緯

医師国家試験出題基準において、基本的臨床知識・技能をより重視し、社会的問題にも配慮した内容とするため、平成5年(第87回試験)から「医学総論」が導入された。また、プライマリ・ケアに関連する医学・医療の基本的事項の出題を容易にするため、平成9年(第91回試験)から「必修の基本的事項」(以下、必修問題)が導入された。その後、約30年が経過する間に、平成13年の医学教育モデル・コア・カリキュラム(以下、コアカリ)策定、平成16年の臨床研修制度開始、平成18年の共用試験(CBT, OSCE)全国実施など、医学教育を取り巻く状況が大きく変化してきた。また、全国各医学部での診療参加型臨床実習の充実が少しずつ

進む中、令和2年の『医師国家試験改善検討部会報告書』(以下、改善検討部会報告書)に示された「医師国家試験の出題基準を可能な限り臨床実習における経験を評価する内容に絞るよう見直しを行うことが妥当である」の提言も踏まえ、医師国家試験においても臨床実習の成果を問う出題内容が重視されている。

2. 必修問題および医学総論問題における出題水準の妥当性検証の必要性

全国医学部長病院長会議(AJMC)が毎年実施している受験生および大学教職員等を対象とした「医師国家試験に関するアンケート調査」の令和5年度(第117回国試)報告では、「臨床実習の成果を問う問題は多かったか」という質問に対し、「多数あった」または「やや多かった」と回答した受験生

の割合は、臨床実地問題で 55.9%、一般問題で 43.6%、必修問題で 46.8%となっている。その一方で、出題者や受験生でない第三者の視点で、医師国家試験の個別問題に対して、臨床実習の成果を問う出題水準となっているかを検証したデータは見られない。また、特に第 116 回国試以降、医学総論・各論の合格基準が必修の絶対基準(80%)に迫る 72~76%台の高い水準で推移しており、出題基準が類似する必修問題と医学総論問題の役割・位置付けやあり方についても今一度検討を行うべき時期に来ている可能性もある。

このような経緯を踏まえ、本研究では、必修および医学総論問題に関する課題抽出と改善への提言を目的として、第 118 回国試の必修および医学総論の個々の問題について、医師養成課程のどの時点の学修成果を問う内容となっているか、第三者的立場の検証パネラーを設定して検証した。

B. 研究方法

1. 検証の対象問題

令和 6 年 2 月に実施された、第 118 回国試の必修問題(100 問)および医学総論問題(150 問)の計 250 問を対象とした。

2. 検証パネラーの選出

検証パネラーの選出要件として、各基本専門領域での診療経験を有し、卒前・卒後医学教育部門の教員として部門管理や学生・研修医指導経験がある医師とし、18 名を選出した。このうち 12 名が日本医学教育学会認定医学教育専門家であり、4 名は本研究班メンバーである。必修問題は 13 名、医学総論問題は 14 名をそれぞれのパネラーとして検証を行った。

3. 検証方法

必修および医学総論問題について、医学部・臨床研修・専門研修のどの時期の学修成果を問う問題か、レベル分類を行った。臨床実習前の学修で回答可能(レベルⅠ)、臨床実習中の学修成果を問う(レベルⅡ)、臨床研修で修得すべき(レベルⅢ)、専門研修で修得すべき(レベルⅣ)の 4 段階を設定した。レベルⅡは、見学を主体とする臨床実習で回答可能な内容をⅡa、診療参加型臨床実習で回答可能な内容をⅡbと区分した(図 1、表 1)。

表1 検証レベルの設定

レベル	概要	具体的内容	レベルに対応する試験
I	臨床実習前の学修で回答可能	基礎医学や公衆衛生学の一般的知識、および各疾患・病態についての一般的な知識や診断を問う内容	共用試験(CBT, Pre-CC OSCE)
II a	臨床実習中の学修成果を問う	疾患・病態の診断および初期対応、救急対応について見学型実習でも修得可能な内容	医師国家試験
II b	臨床実習中の学修成果を問う	疾患・病態の診断および初期対応、救急対応について診療参加型実習の中で修得可能な内容	医師国家試験
III	臨床研修で修得すべきもの	疾患・病態の診断や治療判断を行う上で、一定数の受持患者担当経験や継続的診療経験を必要とする内容	(基本的臨床能力評価試験等)
IV	専門研修で修得すべきもの	疾患・病態の診断や治療判断を行う上で、高度な専門的知識および臨床経験を必要とする内容	各基本領域の専門医試験

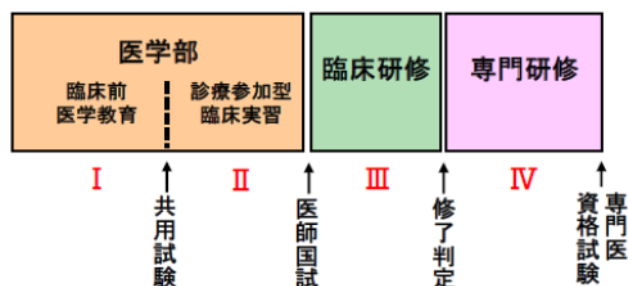


図1 検証レベルの図示

4. 検討項目と統計解析

必修および医学総論問題それぞれについて、事前に 1) 正答率とその分布、2) 正答率と識別指数の関連を解析した。その後、パネラーの検証結果に基づき、3) 各レベル分類の割合、4) レベルⅠ問題の出題割合、5) レベル分類と出題者が設定したタキソノミーの関連、について検討した。

統計解析は、2 群間の割合の比較には、各パラメータの期待度数に応じて、 χ^2 乗検定または Fisher 直接確率法を用い、 $p < 0.05$ の場合を有意差ありとした。

5. 研究の守秘に関して

①研究者の範囲とデータマネジメントプラン(DMP)、②本研究における検証のために、開示レベルに応じて取り扱った研究者、③守秘情報を扱うにあたってのレベルⅤ研究者の実施体制については、研究 3-1 の「B. 方法」に記載したものと同一である。

C. 研究結果

1. 必修問題

1) 正答率とその分布

平均正答率は、全体(100 問)で 90.4%、一般問題(50 問)で 89.7%、臨床実地問題(50 問)で 91.1%であった。正答率の分布は、正答率 95% 以上の問題が60%、正答率 90~94.9%が11%、正答率 80~89.9%が 13%、正答率 70~79.9%が 6%、正答率 70%未満が 10%であった。

2) 正答率と識別指数の関連

正答率と識別指数に一定の逆相関が見られた。正答率 80%未満の問題では識別指数 0.3 以上の問題が約 5 割でみられた。一方、正答率 80% 以上の問題では識別指数 0.2 未満の問題が半数以上を占め、特に正答率 95%以上の問題では識別指数 0.1 未満の問題が 45%でみられた。

3) 各レベル分類の割合

検証パネラーが「臨床実習の成果を問う内容の問題」(レベルⅡ)と判定した設問の割合は、一般問題で 41%(うちⅡa は 21%、Ⅱb は 20%)、臨床実地問題で 69%(うちⅡa は 35%、Ⅱb は 34%)だった。一方、「臨床実習前の学修で回答可能」(レベルⅠ)と判定した設問の割合は、一般問題で 55%、臨床実地問題で 24%だった。

4) レベルⅠ問題の出題割合

個別の問題について、13 名の検証パネラーのうち過半数の 7 名以上がレベルⅠと判定した問題を「レベルⅠ問題」と定義した場合、一般問題では 26 問(52%)、臨床実地問題では 1 問(2%)となり、一般問題が臨床実地問題よりレベルⅠ問題の割合が有意に高かった($p<0.01$)。

5) レベル分類と出題者タキソノミーの関連

教育目標のタキソノミーの考え方にに基づき、個々の問題をⅠ型(想起)、Ⅱ型(解釈)、Ⅲ型(問題解決)に区分した場合、一般問題ではタキソノミーⅠ型の問題が多く出題されていた。一方、臨床実地問題ではタキソノミーⅡ型およびⅢ型の問題が多く出題されていた。

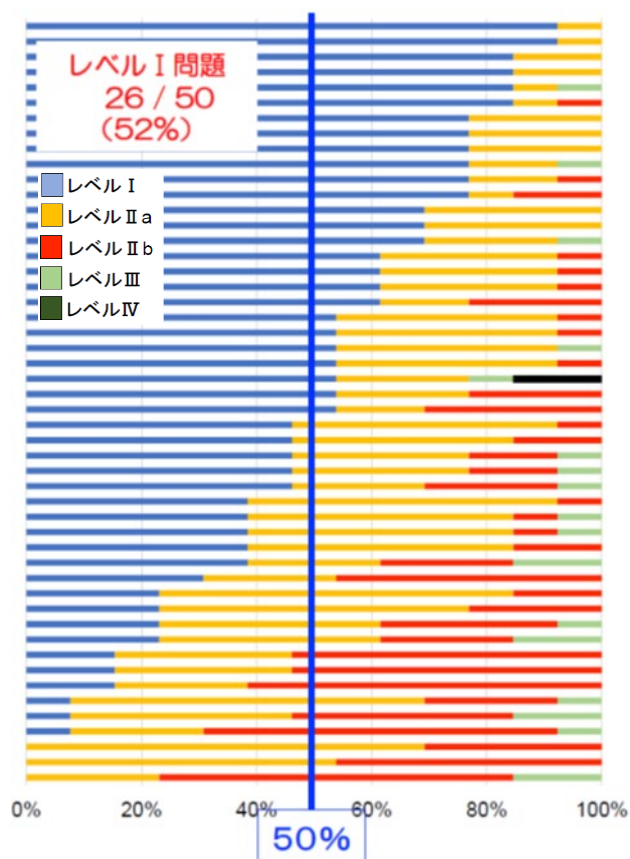


図 2-1 必修(一般問題)におけるレベル分布

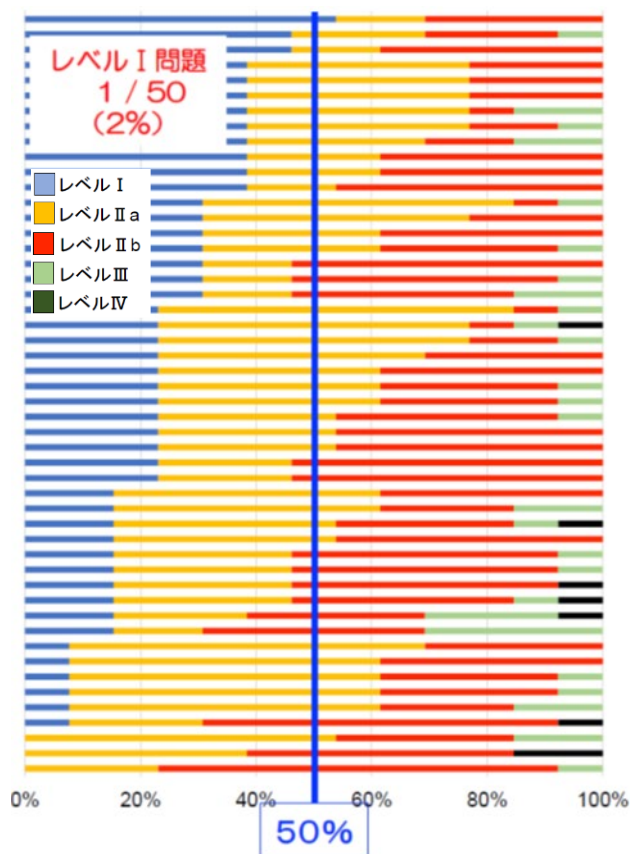


図 2-2 必修(臨床実地問題)におけるレベル分布

2. 医学総論問題

1) 正答率とその分布

平均正答率は、全体(150 問)で 84.5%、一般問題(70 問)で 84.3%、臨床実地問題(80 問)で 84.9%であった。正答率の分布は、正答率 95% 以上の問題が38%、正答率 90~94.9%が21%、正答率 80~89.9%が 13%、正答率 70~79.9%が 9%、正答率 70%未満が 19%であった。

2) 正答率と識別指数の関連

正答率と識別指数は一定の逆相関が見られた。正答率 80%未満の問題では識別指数 0.3 以上の問題が約 3 割でみられた。一方、正答率 90% 以上の問題では識別指数 0.2 未満の問題が約 8 割を占め、特に正答率 95%以上の問題では識別指数 0.1 未満の問題が 32%でみられた。

3) 各レベル分類の割合

検証パネルが「臨床実習の成果を問う内容の問題」(レベルⅡ)と判定した設問の割合は、一般問題で 45%(うちⅡa は 30%、Ⅱb は 15%)、臨床実地問題で 73%(うちⅡa は 34%、Ⅱb は 39%)だった。一方、「臨床実習前の学修で回答可能」(レベルⅠ)と判定した設問の割合は、一般問題で 47%、臨床実地問題で 16%だった。

4) レベルⅠ問題の出題割合

個別の問題について、14 名の検証パネルのうち半数の 7 名以上がレベルⅠと判定した問題を「レベルⅠ問題」と定義した場合、一般問題では 37 問(53%)、臨床実地問題では 5 問(6%)となり、一般問題が臨床実地問題よりレベルⅠ問題の割合が有意に高かった($p<0.01$)。

5) レベルⅠ問題と出題領域の関連

医学総論において、「レベルⅠ問題」とされた設問は、出題基準大項目の一部領域で比較的多く認められる傾向があった。

6) レベル分類と出題者タキソミーの関連

教育目標のタキソミーの考え方に基づき、個々の問題をⅠ型(想起)、Ⅱ型(解釈)、Ⅲ型(問題解決)に区分した場合、一般問題ではタキソミーⅠ型の問題が多く出題されていた。一方、臨床実地問題ではタキソミーⅡ型およびⅢ型の問題が多く出題されていた。

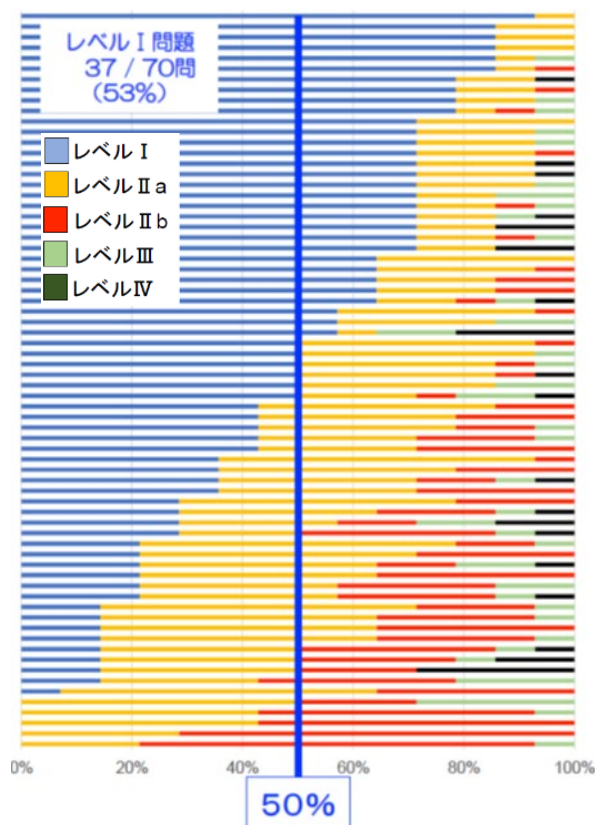


図 3-1 医学総論(一般問題)におけるレベル分布

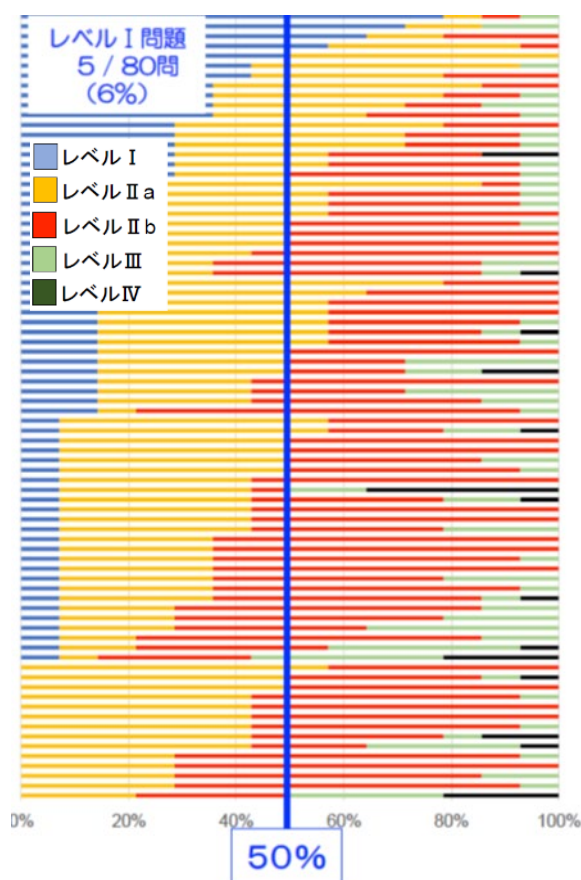


図 3-2 医学総論(臨床実地問題)におけるレベル分布

D. 考察

今回の研究では、第118回国試の必修および医学総論問題について、それぞれ10名以上の検証パネラーが「医師国家試験としてどの時期の学修成果を問う問題であったか」に関してのレベル分類を行い、各問題の医師国家試験での出題妥当性について検証した。その結果、必修および医学総論問題の双方で、一般問題では臨床実地問題より臨床実習の成果を問う問題の割合が少なく、問題内容も想起で回答を得られる、いわゆる解釈や問題解決能力より“浅い”能力を問う問題になりやすいことが示された。また、絶対基準で判定している必修問題と相対基準で判定している医学総論問題が、全体正答率と識別指数やタキソノミー等との関連性において類似している部分が多いことも明らかになった。

図4に近年の医師国家試験の変遷(改善検討部会報告書資料より抜粋)を示す。平成30年からは医師国家試験の一般問題が250問から150問へ削減されたが、その要因として、平成18年から共用試験CBTが全国医学部で開始され、その後も安定的に実施されたことがある。

近年の医師国家試験の変遷				
回	第87～90回	第91～94回	第95～111回	第112回～
年	平成5～8年	平成9～12年	平成13～29年	平成30年～
一般問題	内容 医学総論 医学各論	必修 医学総論 医学各論	必修 医学総論 医学各論	必修 医学総論 医学各論
	数 200問	200問	250問	150問
臨床実地問題	内容 医学総論 医学各論	必修 医学総論 医学各論	必修 医学総論 医学各論	必修 医学総論 医学各論
	数 120問	120問	250問	250問
設問数	計320問		計500問	計400問
試験日数	2日間		3日間	2日間

図4:近年の医師国家試験の変遷

今回の検証結果では、臨床実習前の学修で回答可能な「レベルⅠ問題」の割合が、必修の一般問題50問のうち55%、医学総論の一般問題70問のうち47%を占めており、実に約半数の問題は「臨床実習の成果を問う問題ではない」と判定された。医師国家試験は医師法に基づき厚生労働大臣より任命された出題委員が同一年度の一般問題と臨床実地問題を作成しており、今回の検証でも臨床実地問題では臨床実習の成果を適切に問う問題を多く作成していた。この点を勘案すると、令和2年の改善検討部会報告書に示された「医師国家試験の出題基準を可能な限り臨床実習における経

験を評価する内容に絞るよう見直しを行うことが妥当である」の提言を実行するためには、一般問題という「出題形式」そのものに一定の限界があることを示唆している。改善策としては、現在の医師国家試験は平成30年以前と比較して相対的に臨床実地問題の比重が大きくなっているが、今後は一般問題を更に削減することが考えられる。その際、臨床実習前の共用試験CBTが公的化されていることを考慮し、共用試験CBTとの出題基準の整合性を図ることが必須となる。ただし、これまで、平成28年度版コアカリ(旧版)を出題範囲としてきた共用試験CBTと医師国家試験の出題基準との摺り合わせが十分行われてきたとは言い難い部分もあり、今後は共用試験実施評価機構(CATO)との協議・連携が望まれる。今回の結果を別の視点から捉えれば、一般問題形式であっても約半数は臨床実習の成果を問う問題が作成されていることになる。したがって、今回の研究結果をさらに詳細に検証し、「どのような要素があれば臨床実習の成果を問う一般問題が作成できるのか」を抽出し、出題委員へ還元できれば、問題の質の改善が期待される。また、医師国家試験がCBT化され、動画や音声・連続画像などが試験で出題できるようになれば、一般問題形式であってもより臨床実習の成果を問う問題を作成しやすくなることが期待される。

必修と医学総論問題の相似性については、平成5年の医学総論導入から4年後に必修問題が導入された時系列と、必修と医学総論の出題基準の重複(必修の基本的事項18項目のうち14項目は医学総論との重複がある)を鑑みれば、ある程度想定される事象と思われる。この20年で継続的に改訂されているコアカリが各大学の医学教育カリキュラムへ取り込まれていること、共用試験が安定的に実施されて公的化したこと、医学教育分野別認証評価を全医学校が受審していることなどから、各大学医学部の教育は標準化と均てん化が進み、医学生が修得している知識や技能の質も向上している。これらを総合的に勘案すると、必修と医学総論の双方が必要であるかについての議論を始める時期である。

同分担研究「令和6年度版医師国家試験出題基準改定に伴う第118回医師国家試験の検証」におけるインタビュー報告にも記載されているとおり、医師国家試験の出題委員が医学総論問題と必修問題の作問意図や難易度を厳密に区別して作問することが技術的に難しくなっているという指摘がある。この点については、本研究の研究デザイ

ンと結果から言及することが難しい部分もあり、出題者への更なるインタビュー調査など、今後の更なる検討が期待される。

本研究の limitation としては、検証パネラーの選択の適切性、また、研究方法として設定した学修時期に基づくレベル分類の妥当性などが考えられる。今回の検証パネラーは、卒前・卒後医学教育部門の責任者あるいは指導者として医学教育全般を俯瞰的に見ることでできる臨床医学系の教育エキスパートの方々を選出した。また、選出に際しては、内科、外科などの診療経験領域も極力重ならず、かつ小児科、精神科など幅広く網羅できることを意識した。しかしながら、今回の検証パネラーの中には耳鼻科や眼科等の全ての領域の専門医取得者を含んではいなかったため、今後は、より多様な視点から現在の国家試験の出題について検証を得られる研究スキームが必要と考える。また、分類レベルについては、主に「適切な学修成果を問う時期」に重点を置いて設定し、表 1 に示す分類条件以外は検証パネラーの判断に委ねた。このレベル分類についても、今後の更なる検討が可能かもしれない。

【今後への提言】

- 平成 9 年導入の必修問題に関して、コアカリ、臨床研修、共用試験などの教育改革を経て今日的な意義を再検討し、医学総論との発展的な統合化を含めた根本的な見直し強く望まれる。
- 医学総論での社会医学系、基礎医学系および臨床病態生理領域等、臨床実習前の CBT 等で問うべき内容とレベルに関して、医療系大学間共用試験実施機構(CATO)との協議を行うべきである。
- 資質・能力に基づいたコアカリの記載と医師国家試験出題基準との整合を強く推奨するとともに、必修および医学総論では、診療参加型臨床実習の成果を問うような試験の枠組(ブループリント)の見直しについても検討することが望まれる。

E. 結論

今回の妥当性検証の研究により、大きく以下の 4 つの結果が得られた。

1)『臨床実習中の学修成果を問う問題レベル(Ⅱa,Ⅱb)』と判断した割合は、必修では一般問題 41%、臨床実地問題 69%、医学総論では一般問

題 45%、臨床実地問題 73%で、臨床実地問題形式の方が一般問題より臨床実習中の学修成果をより問える割合が高かった。

2)『臨床実習前の学修で回答可能な問題レベル(Ⅰ)』と判断した割合は、必修では一般問題 55%、臨床実地問題 24%、医学総論では一般問題 47%、臨床実地問題 16%で、一般問題形式の方が臨床実地問題より臨床実習前の学修で解答可能な内容を問う割合が高かった。

3)パネラーの半数以上がレベルⅠ問題と判定した問題の割合は、必修・医学総論双方で一般問題が臨床実地問題より有意に高かった(必修:52% vs. 2%, $p < 0.01$ 、医学総論:53% vs. 6%, $p < 0.01$)。

4)医学総論では、レベルⅠ問題とされる設問が、出題基準の一部領域で比較的多く認められ、領域によっては共用試験との出題範囲の調整を要する可能性が示唆された。

F. 健康危険情報

本研究で該当する事項はない。

G. 研究発表

1. 論文発表

1)日本医学教育学会雑誌「医学教育」に投稿予定である。

2. 学会発表

1)報告書記載時点で学会発表予定はない。

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

磯部 真倫	(岐阜大学 産婦人科)
大久保 由美子	(帝京大学 医学部 医学教育学講座)
近藤 猛	(名古屋大学医学部附属病院 総合診療科/ 卒後臨床研修・キャリア形成支援センター)
高橋 誠	(北海道大学 大学院医学研究院 医学教育・ 国際交流推進センター)
高見 秀樹	(名古屋大学医学部附属病院 卒後臨床研修 キャリア形成支援センター)
永田 康浩	(長崎大学 生命医科学域(医学系))
西屋 克己	(関西医科大学 教育センター)
野村 理	(岐阜大学 医学教育開発センター)
蓮沼 直子	(広島大学 医学部附属医学教育センター)
藤田 博一	(高知大学 医学部附属医学教育創造センター)
松島 加代子	(長崎大学病院 医療教育開発センター)
松山 泰	(自治医科大学 医学教育センター)
三井 良之	(近畿大学 医学部教育センター)
山本 憲	(順天堂大学 健康データサイエンス学部)

