

医師国家試験出題基準に関する研究

～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～

研究分担者 小西 靖彦(順天堂大学医学教育研究室 特任教授)
小松 弘幸(宮崎大学医学部 医療人育成推進センター 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

令和2年の医師国家試験改善検討部会報告書の内容を受け、令和6年度版医師国家試験出題基準への改定が行われた。この基準に沿って初めて行われた第118回医師国家試験を検証した。

最大の改定点である、掲載疾患・病態の頻度・重要度に応じた医学各論での「病名の a,b,c 分類」の導入により、第118回医師国家試験において a 病名からの出題が第117回と比べて 8.6 ポイント増加(64.7→73.3%)した。医師国家試験委員に対するインタビューとアンケートからも、a,b,c 分類は試験問題作成プロセスに有効に機能していることが確認された。回答肢においては、a,b,c 病名に該当しない 29 病名(全回答肢病名の 9.8%)や備考欄にある 37 病名(同 12.5%)がある等の課題が認められた。

出題基準医学各論の備考欄には 476 項目が記載され、うち 283 項目(59.5%)が a,b,c の適用を受けない病名であった。第116回～第118回医師国家試験で、それぞれ備考欄から 58,71,71 題が出題されており、殆どはこれら病名からであった。

これらの事実から、医学各論の小項目には「病名」を中心とした記載に統一することを原則とし、備考欄の整備を含めて出題領域間で整合をとる必要がある。医学各論の出題基準がどの疾患についてどの深さで設問するかのガイドライン機能に限定する等、基準の位置付けに見直しが必要であると考えられた。

A. 研究目的

国は、「医師国家試験として妥当な範囲と適切なレベルを保ち、医師の資質の向上を図る」ために、定期的に「医師国家試験改善検討部会」(以下、改善検討部会)を開催して、現行の医師国家試験を評価し改善について検討してきた。直近の改善検討部会は、令和2年2月に報告書をまとめ、さまざまな改善事項を挙げている。これを受けて「令和6年版医師国家試験出題基準」(以下、R6 出題基準)が改定された。最大の改定点は、医学各論における掲載疾患・病態の頻度・重要度に応じたレベル分類(以下、病名の a,b,c 分類)の導入であった。

この R6 出題基準を受けて初めて出題が行われ

た第118回医師国家試験(令和6年2月施行)の状況を検討し、その効果について検証すること、および、第118回医師国家試験の出題プロセスへの R6 出題基準改定の影響を検証することが本研究の主目的である。

B. 研究方法(第118回国試験検証)

1. 第118回医師国家試験の検証に関して

R6 出題基準における最大の改定点は、「医学各論」における「病名の a,b,c 分類」の導入であった。病名の a,b,c 分類について第118回の試験内容を次の方法で検証した。

1)第 118 回国試「医学各論」における a,b,c 病名の出題数の検討

- ・ R6 出題基準における医学各論の大項目別 a,b,c 病名の実数とその分布を調査した。
- ・ 第 118 回国試の医学各論問題(150 題)における問題内容の a,b,c 病名別の実数を医学各論の大項目(Ⅰ～XIII)別にピックアップした。
- ・ 比較を目的として、第 117 回国試の医学各論(150 題)で扱われた問題内容の a,b,c 病名別の実数を後方視的に判断し、大項目(Ⅰ～XIII)別に行った。

2)第 118 回国試「医学各論」における a,b,c 病名群別の正答率の検討

- ・ a,b,c 病名群別に正答率の分布および平均正答率を算出した。

3)第 118 回国試「医学各論」について、回答肢として採用された病名が a,b,c 病名のいずれに該当するか、あるいは該当しないのかについて全 150 問の検討を加えた。

- ・ 後方視的な検証である。
- ・ 回答肢の病名が、医学各論の小項目等だけでなく備考欄からも採用されていたため、これらもカウントした。
- ・ 回答肢における a,b,c 病名の出現数と率に加えて、a,b,c 病名に含まれない回答肢病名および備考欄に記載のある回答肢病名の数と率を検証した。

4)R6 出題基準における「医学各論備考欄」に関する検証

- ・ R6 出題基準の「医学各論」において備考欄の全ての記載事項を抽出し、類型化を行った。
- ・ 医学各論の大項目(Ⅰ～XIII)別に、備考欄の掲載数を調べてカテゴリーに類型化した。カテゴリーは[追加病名][検査・術式・治療等][説明・付加症状]を設定した。

5)「医学各論備考欄」から出題されていた状況およびその経時的変化に関する検証

- ・ 第 116 回から第 118 回の 3 年間の医学各論問題について、R6 出題基準の医学各論備考欄にある記載項目から出題が行われていたかについて検証を行った。
- ・ 第 116,117 回国試においては、R6 以前の出題

基準が使用されていたため、比較を目的として R6 出題基準の備考欄記載項目を基準に各回の問題を検証した。

- ・ 医学各論の大項目(Ⅰ～XIII)別に、3 年間の出題数とその変化について検証を行った。

6)医学各論において病名の a,b,c 分類が 118 回国試の問題作成委員に与えた影響分析

R6 出題基準における“病名の a,b,c 分類”の導入が、医師国家試験の作成にあたって出題委員にどのような影響を与えたかの検証を行った。

(1)中心的な出題委員への影響(インタビュー調査)

- ・ 第 118 回出題委員のなかで経験年数が比較的長く、出題会議で全体を見回す役割を負ったと考えられる 4 名を選んだ。インタビューの選定は研究者 2 名で行い、専門の領域が重ならないように配慮した。
- ・ 4 名の背景(概要)は次のとおりである。必修を主に担当する者 1 名、医学総論を主に担当する者 1 名、医学各論を主に担当する者 2 名(それぞれに担当の重複はある)。
- ・ インタビュー時期は令和 6 年 7 月から 8 月にかけて、Web を利用した Semi-structured Interview とした。
- ・ 研究者 2 名がインタビュー 1 人当たり 60 分程度の質問を行った。
- ・ 質問内容の概要は事前に送付した。その内容は次のとおりである。

- ① 問題(個人分担)作成時に、R2 改善検討部会報告書の内容はどの程度意識していたか？
- ② 問題(個人分担)作成時に、R6 出題基準の医学各論での病名の a,b,c 分類はどの程度意識していたか？
- ③ 医学各論での病名の a,b,c 分類が示されたことで、自身の作成問題の疾患選定や難易度に影響があったと思うか？
- ④ 医学各論での病名の a,b,c 分類が示されたことで、出題委員会での審議内容に影響があったと思うか？
- ⑤ 医学各論での病名の a,b,c 分類についての個人的な自由意見
- ⑥ 第 118 回国試全体についての個人的な自由意見

(2) 主要な出題委員に対するアンケート調査

- ・ (1)のインタビュー結果を参考として、主要な出題委員に対するアンケート調査を行った。
- ・ 医師国家試験の作成に関して、ブラッシュアップ等の主要な役割を担うメンバーに対するものであるが、その氏名と背景および総人数などは公表されていないため、本報告書では詳細を記載しない。
- ・ アンケートは令和6年10月から11月にかけてWebを利用して行った。
- ・ 数十名の主要な委員全員から回答を得た。
- ・ 質問項目は次のとおりである。

- ① 問題作成時に病名の a,b,c 分類が与えた影響
 - ・ 問題作成時の医師国家試験改善検討部会報告書への意識
 - ・ 問題作成時のレベル分類への意識
 - ・ レベル分類が疾患選定、難易度に与えた影響
 - ・ グループ内でのレベル分類に対する異論、不満
- ② 試験問題のブラッシュアップに 病名の a,b,c 分類が与えた影響
 - ・ レベル分類の影響
 - ・ 領域間でのレベル選択基準の差
- ③ 第118回国試全般に関する意見
 - ・ 試験全体の難易度
 - ・ 備考欄への考え
 - ・ 第118回の禁忌肢への考え
 - ・ 第118回の英語問題への考え
 - ・ 第118回の計算問題への考え
- ④ 医学総論・必修に関する意見
 - ・ 医学総論の臨床実地問題が特定の疾患、病態に偏っているとする指摘への考え
 - ・ 必修と医学総論の大～中項目との重複について
 - ・ 必修の位置づけへの考え
 - ・ 医学総論の位置づけへの考え
 - ・ 医学総論での保健医療/公衆衛生について(問題数および難易度)
- ⑤ その他自由意見

7) 不適切問題、禁忌肢問題に関する検証

主研究者である小西靖彦と小松弘幸は、厚生労働省医事課試験免許室に誓約書を提出し、過去に遡った不適切問題および禁忌肢問題に関する情

報にアクセスした。

- ・ 妥当性等に関する検証を行い、研究結果は、試験免許室と情報共有した。
- ・ この点に関する研究結果は、国家試験に関する公表されていない情報に基づいており、本研究報告書には記載しない。

2. 研究の守秘に関して

試験問題の検証にあたっては厚生労働省医政局医事課試験免許室の協力を得て、各回の問題情報を得た。

1) 守秘すべき情報の扱い

(1) 研究者の範囲とデータマネジメントプラン(DMP)

データの重要性和秘匿性、および個人情報保護の観点に基づいて機密性レベルを5段階に分類し、開示レベルとデータに触れることができる者を次のように設定した。

表1 機密性レベルの5段階分類

機密性レベル	開示レベルの定義	データに触れることができる想定対象者
レベルⅠ	医師国家試験に関する秘匿性が低く本研究報告書に記載されているレベルに相当するもので、個人を同定できる情報を含まないレベル	・ 本研究報告書の読者 ・ 本研究が掲載される厚生労働科学研究成果データベースにアクセスできる一般の方
レベルⅡ	医師国家試験に関する数値を含む具体的データを含まないもの及び必修・医学総論・医学各論に関する検証の要約等に触れるレベルであり、個人を同定できる情報を含まないレベル	・ 本研究においてインタビュー／アンケート研究に協力した医師国家試験の委員、及び「医学総論」「必修の基本的事項」の検証パネラーとして参加した者
レベルⅢ	上記開示レベルⅡに加え、各論検証における備考欄および主要な国家試験委員へのインタビュー／アンケート内容、必修・医学総論・医学各論に関する検証の要約等に触れるレベルであり、個人を同定できる情報を含まないレベル	・ 厚労科研河北班の全体研究に参加する者
レベルⅣ	上記開示レベルⅢに加え、不適切問題・禁忌肢の検証、医学総論・必修におけるパネラー検証のデータ全体に触れることのできるレベルであり、個人を同定できる情報を含まないレベル	・ 厚生労働省試験免許室の室員 ・ 試験免許室に対して守秘情報に関する利用誓約書を提出してデータにアクセスすることが許された試験に関する専門家
レベルⅤ	本研究におけるすべてのデータを管理し、主要な試験委員に対するインタビュー／アンケート結果及びパネリストからの回答等、個人を同定できる情報を扱うレベル	・ 本研究の主研究者(小西靖彦、小松弘幸)

(2) 本研究における正確な検証のために、公表されている試験問題、正答肢などだけでなく次のような情報を得て研究を進めた。これらの情報は開示レベルⅣとⅤにある者に限定した。

- ・各問の正答率、識別指数
- ・出題者が作問時に想定した出題意図、病名のa,b,c分類、タキソノミー等
- ・いわゆる不適切問題とされた設問の正答率、識別指数等
- ・出題された禁忌肢の内容および受験者の当該肢の選択状況等
- ・その他、医師国家試験で非公開となっている情報

『医学各論』のa,b,cの出題数統計

【出題科目】 大項目	小項目	【出題された病名数】				【出題された禁忌肢数】				【出題された非禁忌肢数】			
		a	b	c	合計	a	b	c	合計	a	b	c	合計
1. 内科学	内科学全般	100	20	10	130	4	6	3	13	4	2	2	8
2. 外科	外科全般	120	25	15	160	3	2	1	6	3	1	1	5
3. 小児科	小児科全般	80	15	5	100	2	1	1	4	2	1	1	4
4. 産婦人科	産婦人科全般	110	20	10	140	3	2	1	6	3	1	1	5
5. 皮膚科	皮膚科全般	60	10	5	75	1	1	1	3	1	1	1	3
6. 泌尿器科	泌尿器科全般	70	12	6	88	2	1	1	4	2	1	1	4
7. 耳鼻咽喉科	耳鼻咽喉科全般	50	8	4	62	1	1	1	3	1	1	1	3
8. 眼科	眼科全般	40	6	3	49	1	1	1	3	1	1	1	3
9. 放射線科	放射線科全般	30	4	2	36	1	1	1	3	1	1	1	3
10. 精神科	精神科全般	20	3	1	24	1	1	1	3	1	1	1	3
11. 法医学	法医学全般	10	2	1	13	1	1	1	3	1	1	1	3
12. その他	その他	10	2	1	13	1	1	1	3	1	1	1	3
合計		710	140	70	920	25	25	15	65	25	15	10	50

図1 医学各論におけるa,b,c分類毎の出題数

(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

- ・小西、小松の両名は本文分担研究に関するデータマネジメントプラン(DMP)を作成し、それに基づいた情報の取り扱いを行った(概要を①に記載した)。
- ・両名は、本研究の実施にあたり厚生労働省の試験免許室に対して、医師国家試験情報の利用に係る誓約書を提出して作業を行った。
- ・両名は、過去に医師国家試験の主要な委員として活動した経験があり、その際も誓約書を提出して作業を行っていた。
- ・両名は、それぞれの所属機関で研究倫理教育を受講し、COI委員会に申し出を行っている。

C. 研究結果(第118回国試験検証)

1. 第118回国試験「医学各論」におけるa,b,c分類毎の出題数の検討

- ・R6出題基準に掲載された医学各論のa,b,c分類は、a病名が643(43.4%)、b病名が537(36.2%)、c病名が302(20.4%)であった(図1)。
- ・大項目ごとのそれぞれの病名数(うち数)は図1に示すとおりである。
- ・第118回の医学各論における病名a,b,c分類の分布は、a病名が110(73.3%)、b病名が32(21.3%)、c病名が5(3.3%)であった。
- ・第117回の医学各論における病名a,b,c分類の分布は、a病名が97(64.7%)、b病名が37(24.7%)、c病名が11(7.3%)であり、R6出題基準の改定を経て、a病名が8.6ポイント増加し、c病名が4.0ポイント減少した。

2. 第118回国試験「医学各論」における病名a,b,c分類別の正答率の検討

- ・第118回国試験の医学各論全体150問の正答率の平均は85.0%であった。
- ・このうち、a,b,c病名別の平均正答率は、a病名(110題)が86.4%、b病名(32題)が84.0%、c病名(5題)が74.3%であった。

3. 第118回国試験「医学各論」の回答肢におけるa,b,c病名への該当/非該当

- ・第118回国試験医学各論の設問における回答肢に採用された病名について検討を加えた(図2)。医学各論の回答肢には延べ296病名が採用されていた。その内訳は、a病名が142(47.9%)、b病名が68(23.0%)、c病名が20(6.8%)であった。
- ・それ以外に、「回答肢がa,b,cどれにも該当しない病名」が29(9.8%)、「回答肢が医学各論の備考欄から出題された病名」が37(12.5%)存在した。
- ・それぞれ、正答肢として出題されたか誤答肢として出題されたかの内訳を図2に示すとおりで、「回答肢がa,b,cどれにも該当しない29の病名」のうち6(21%)が正答肢、23(79%)は誤答肢として出題されていた。「回答肢が医学各論の備考欄から出題された37の病名」のうち17(46%)が正答肢、20(54%)は誤答肢として出題されていた。
- ・「回答肢がa,b,cどれにも該当しない病名」および「回答肢が備考欄から出題された病名」の具体的な病名については図3に示す。

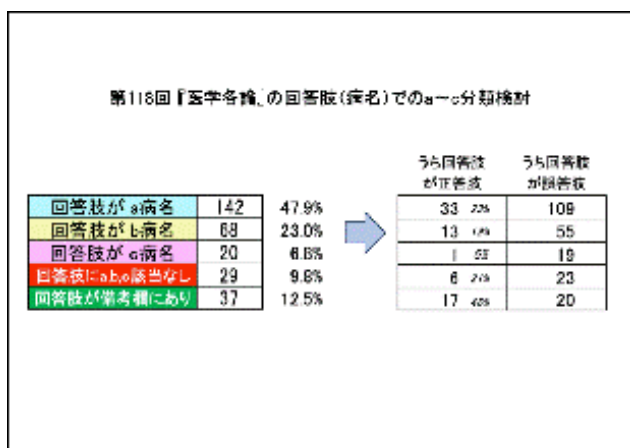


図2 第118回国試医学各論の回答肢のa,b,c分類



図3 第118回国試医学各論でのa,b,c病名以外からの回答肢

4. R6 出題基準における医学各論備考欄に関する検証

- ・ R6 出題基準の医学各論における備考欄には、330箇所(箇所の備考欄に延べ476)の記載があった(図4)。
- ・ I から XIII までの医学各論の大項目別にみると、最低8から最大80まで記載数の差がみられた。
- ・ 備考欄の記載を[追加病名][検査・術式・治療等][説明・付加症状]の3つのカテゴリーに分けて類型化を行ったところ、備考欄に記載された476項目のうち283項目(59.5%)が追加病名であり、検査・術式・治療等は62項目、説明・付加症状は131項目であった。
- ・ 追加病名を多く記載する大項目領域には、VI 消化器・腹壁・腹膜疾患領域:57項目、IX 神経・運動器疾患領域:46項目、X 内分泌・代

謝・栄養・乳腺疾患領域:38項目などがある一方、追加病名をあまり記載していない領域(II 精神・心身医学的疾患:5項目)などがある(図4)。

- ・ 内分泌疾患の一部などで、疾患名を小項目に記載せずに備考欄に記載しているものが認められた(考察の項で詳述する)。
- ・ 検査・術式・治療等の備考欄記載は、特定の領域に多い傾向があった。その領域は、V 心臓・脈管疾患:19項目、III 皮膚・頭頸部疾患:10項目、I 先天異常、周産期異常、成長・発達異常領域:9項目などである。ただし、I 領域では医学総論で記載されている事項が医学各論備考欄に再掲されているものがあった(考察の項で詳述する)。
- ・ 心臓および頭頸部領域において術式の記載が多い傾向があった。
- ・ 付加説明は、領域によって備考欄の使用理由がさまざまであり、その内容と深さに一定性が認められなかった。



図4 第118回国試医学各論の備考欄のカテゴリー別頻度

5. 医学各論備考欄からの出題状況および過去3年間の変化

- ・ 第116回から118回までの3年間の国試医学各論問題を後方視的に解析したところ、第116回では備考欄から58問、第117回で71問、第118回では71問の出題がなされていた。
- ・ 備考欄からの出題には、備考欄の記載事項そのものを主題として問う設問、備考欄記載事項を正答肢あるいは誤答肢とした設問などの類型

があった。

- ・ 備考欄からの出題は追加病名からのものが中心であった。誤答肢としての出題が多い傾向があったが、備考欄の記載事項(病名)そのものを
出題するものや正答肢として問うものもある
(第 118 回国試に関しては、図 3 を参照)。
- ・ 検査・術式・治療等からの出題は多くない。胎
児心拍数陣痛図や Apgar スコアなどが医学各
論備考欄から出題されているが、これらはほと
んど医学総論に記載されている事項であり、医
学各論備考欄に重複記載されたものであって、
これらを除けばこのカテゴリーからの出題は少
ない。症状・付加説明からの説明もほぼ認めず、
備考欄からの出題は概ね「追加病名」のカテ
グリーからであった。

『医学各論』の備考欄からの出題状況									
『医学各論』 大項目	出題 割合	第118回国試の出題率			第117回国試の出題率			第116回国試の出題率	
		【追加病名】 延べ47.6の備考欄記載	【検査・術式・治療等】 延べ1.5の備考欄記載	【症状・付加説明】 延べ0.1の備考欄記載	【追加病名】 延べ47.6の備考欄記載	【検査・術式・治療等】 延べ1.5の備考欄記載	【症状・付加説明】 延べ0.1の備考欄記載	【追加病名】 延べ47.6の備考欄記載	【検査・術式・治療等】 延べ1.5の備考欄記載
先天異常、先天性異常、成長・発達異常	約 1%	10	9	5	7	14	12		
精神・心身医学的疾患	約 1%	5	0	3	2	2	0		
皮膚病・泌尿器疾患	約 11%	23	10	18	8	4	4		
呼吸器・胸壁・縦隔疾患	約 7%	15	1	21	3	3	0		
心臓・血管疾患	約 10%	13	19	11	5	7	9		
消化器・胆道・膵臓疾患	約 15%	57	6	17	14	6	3		
内分泌・代謝疾患	約 1%	24	6	13	7	10	1		
腎臓・泌尿器疾患	約 12%	19	4	6	8	4	7		
神経・運動器疾患	約 9%	46	4	9	7	4	7		
免疫・アレルギー疾患・感染疾患	約 8%	38	2	9	5	10	7		
血液・造血系疾患、膠原病、免疫病	約 2%	12	0	2	0	1	0		
遺伝性疾患	約 8%	20	0	3	4	5	1		
消化・循環器系・免疫性因子による疾患	約 1%	1	1	14	1	1	1		
【計】	100%	283	62	131	71	71	58		

図 5 第 118 回国試医学各論での備考欄からの出題状況と 3 年間の推移

6. 中心的な出題委員へのインタビュー結果

[a,b,c 分類が与えた影響について](インタビューでの主な発言を記載する)

- ・ 研修開始時に必要な知識を問うという方向性、医学各論 a～c 分類はかなり意識した。
- ・ 出題委員の意識は(当初)低かった。
- ・ 主要な委員のほとんどが医学各論病名の a,b,c 分類を意識していた。
- ・ 新規の委員は病名の a,b,c 分類をあまり認識できていなかった印象だが、複数年経験委員の多くは意識して作問してきていた印象である。
- ・ 医学各論問題作成時の難易度設定には、レベル分類を意識したので、影響はあった。

- ・ 委員による出題基準逸脱や難易度の高い問題に対して、レベル分類を根拠にコントロールができた。
- ・ 複数の他領域委員がレベル分類を根拠に指摘してくれて(出題内容が)変更となった。

[118 回国試全体に関する意見](インタビューでの主な発言を記載する)

- ・ 合否水準が 76.7%まで上昇し簡単になりすぎた。
- ・ レベル分類によって、卒業までに学修すべき疾患に出題傾向が寄ってきていると感じる。
- ・ 臨床現場に出しても大丈夫かをみる試験であるべきで、今の方向性で良いと思う。
- ・ 丸暗記だけでは解けない良問が増えたと感じている。
- ・ 易化傾向が続き全体正答率が高くなると識別指数に影響がないか懸念する。
- ・ 予備校などでは英語問題は捨て問題として解かなくていいという話もあると聞いている。
- ・ 公衆衛生は範囲が広く、大学による教育の差も大きいのではと懸念する。
- ・ 保健医療/公衆衛生の領域で備考欄からの出題もあり、問題に使用される図表など全医学生が見ているのかと気にはなった。
- ・ 保健医療/公衆衛生の領域では専門家同士の議論が中心となり、他領域の委員は意見を出しづらい印象がある。
- ・ 多くの委員が必修の必要性・意義が不明と言っていた。
- ・ 必修と医学総論はかなりオーバーラップしており、違いは難易度のみに思える。
- ・ 国試が CBT 化されれば、必修問題は不要になると思う。
- ・ 毎年新規に 400 問作るのは大変で類似問題となりやすく、CBT 化(問題非公開、プール化)が望ましい。
- ・ CBT 化すれば、現在のようにチャンピオン画像 1 枚ではなく、動画や CT の連続画像が出題できる。

7. 主要な出題委員に対するアンケート結果

巻末にアンケート全体像を示し、ここでは主なアンケート結果について記載する。

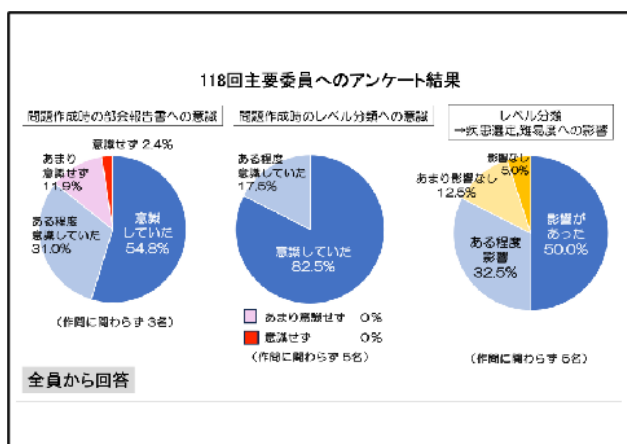


図 6: 第 118 回国試委員への影響に関する回答

- ・ 図 6 に示すとおり、「第 118 回国試問題の作成時に改善検討部会(R2)の報告書を意識していた」と回答した者は、ある程度の意識まで含めると 85.8%であった。
- ・ 「問題作成時に「レベル分類」を意識したか」と問いを変えると意識していたが 82.5%で、ある程度までの意識まで含めると 100%であった。
- ・ 「レベル分類が疾患選定や難易度に影響を与えたか」の問いに対してはある程度までを含めると 82.5%に影響があったと回答された。

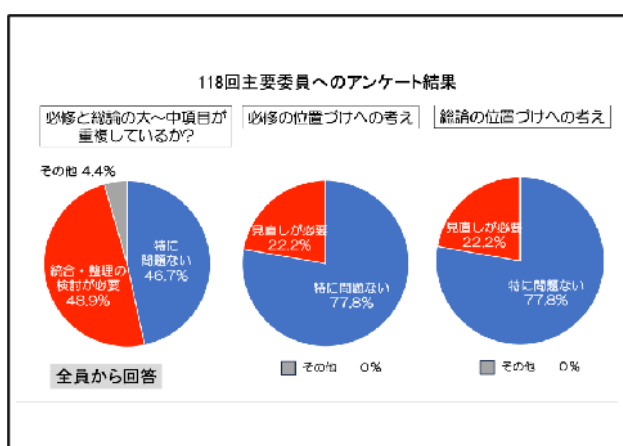


図 7: 必修と医学総論に関する主要委員の回答

- ・ 主要な出題委員に対し「必修と医学総論の重複」について質問した。46.7%が特に問題ないと回答する一方で、48.9%が統合・整理など検

討が必要と回答した(図 7)。特に問題ないとした意見(抜粋)では「基本的事項、重要事項は何回も問うてよい」「医学総論は必修内容でより高難易度の内容を問うため、重複は許容できる」などがあった。一方で統合・整理が必要とする意見(抜粋)では「出題基準の構成に改善の余地がある」「CBT 化により必修問題を作成する必要がなくなるのではないか」「必修・医学総論の名称から重複は当然で、棲み分けは必要」などがあった。

- ・ 必修の位置付けに関しては、特に問題ないが 77.8%で見直しが必要とするものが 22.2%であった。特に問題ないとした意見には「医師にとって身につけておくべき重要な内容である」「位置付けに問題は感じないが、高得点化しすぎていると思う」「医師でなくても回答できるような常識問題は削除する方針が良い」があり、問題ないとしつつ課題が指摘されていた。必修の見直しが必要とする意見(抜粋)には「易しすぎて問題の体をなさないこともあり、医学総論との共存の検討を」「ほぼ全員が正答する試験を作る意味合いを問う」「一般教養的事項 A は削除を」「(共用試験の)CBT 開始から年月も経ち、必修の必要性について議論が必要 →医学総論と医学各論のみとした方が分かりやすい」「高い正答率を求められ道徳的に回答できる極端な易問が多く出題されている」などがあった。
- ・ 医学総論の位置付けに関しては、特に問題ないが 77.8%で見直しが必要とするものが 22.2%であった(必修の位置付けに関する分布と同率であった)。特に問題ないとした意見には「位置づけは良いが、内容の重複やそもそもの項目の修正は必要」があり、問題ないとしながら課題が指摘されていた。医学総論の見直しが必要とする意見(抜粋)には「必修との差が分かりにくい」「公衆衛生的な内容が多すぎるように感じる」「医学各論との違いが難しく、タキソミーⅡ,Ⅲでは治療等となり診療科をまたぐより医学各論疾患に対する対応を聞きがちになる」「Ⅱ-11-C 産業医と労働安全衛生管理の出題数は減らしたり難度を下げたりしても良い」「病院での医師の振る舞い、社会人としての常識を

問う問題の出題を多くしたい」などがあった。

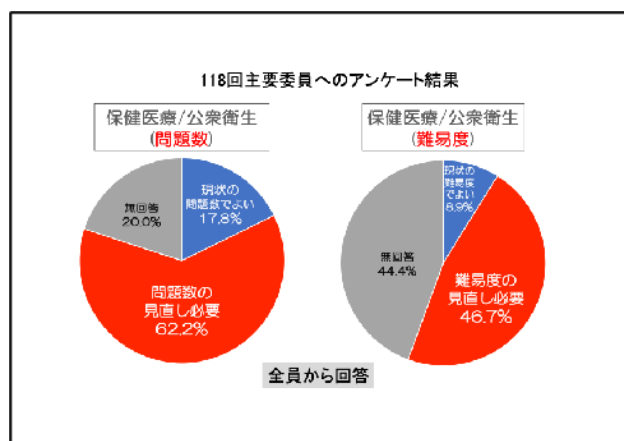


図 8:保健医療/公衆衛生に関する主要委員の回答

- ・医学総論における保健医療/公衆衛生の領域からの出題に関し、「問題数」について尋ねたところ 62.2%から見直しが必要とされ、現状の問題数でよいと回答したものは 17.8%であった(図 8)。問題数に関して特に問題ないとした意見は「患者や同僚と接するうえでの法令、社会的常識については必要」というものであった。問題数に関して見直しが必要とする意見(抜粋)には「研修医が必要な知識から逸脱しており、最低限の知識のみで良い」「臨床系は「臨床研修医に必要なか」が問われるが、保健医療/公衆衛生は必要のない知識を問うことが多く、問題数の大幅な減少を」「社会問題は変動するので減らして良い」「将来の自己研鑽に役立つ基礎医学分野を重視し増やしてほしい」などがあった。
- ・医学総論における保健医療/公衆衛生の領域からの出題に関し、「難易度」について尋ねたところ 46.7%から見直しが必要とされ、現状の難易度でよいと回答したものは 8.9%であった(図 8)。難易度に関して特に問題ないとした意見には「患者や同僚と接するうえでの法令、社会的常識については必要」というものがある一方「重箱の隅の制度問題は不要」とする否定的意見もあった。難易度に関して見直しが必要とする意見(抜粋)には「研修医レベルの実臨床に必要な内容に絞るべきと考える」「法律、制度は臨床現場で役立つ知識という視点で難易度を

調整する必要」「問題と直接結びつかない出題基準が多数あり、公衆衛生の用語集のようになっている印象→整理が必要」「一般医師が解けない資格試験として相応しくないレベルである」などがあった。

D. 考察(第 118 回国試験検証)

1. 第 118 回国試「医学各論」における病名の a,b,c 分類を前提とした出題に関して

- ・病名の a,b,c 分類は良い方向に機能したと考えられる。第 117 回国試と比較した場合、R6 出題基準の改定を経て第 118 回国試では a 病名が 8.6 ポイント増加し、c 病名が 4.0 ポイント減少している。病名の a,b,c 分類の導入が、出題委員に対して難易度の高い病名を避ける方向に働いたと推察される。このことは、結果 C.7) に記載した主要な委員のアンケート(図 6)からも傍証されている。
- ・図 1 に示されたとおり、R6 出題基準に掲載された a,b,c 病名の比率がそれぞれ 43.4%, 36.2%, 20.4%だったところ、第 118 回で出題された比率が a 病名 73.3%、b 病名 21.3%、c 病名 3.3%だった事実は、出題基準にある c 分類の病名等が本当に必要なのかを問いかけている。ただし、作問には回答肢の用意が必要であることに留意しなくてはならず、出題基準が一定程度括めに記載されることは已むを得ない部分もあるだろう。特に、5 つの回答肢を病名で揃えようとする際には、a 病名だけで選択肢を揃えることが難しいことが容易に想像できる。病名の a,b,c 分類が出題の主題となるキーワードなのか、回答肢における出現の制限を伴うものか、意味付けの確認は重要である。なお、本研究では、回答肢における a,b,c 病名について検討を加えたので後述する。
- ・第 118 回国試の医学各論全体 150 問の正答率の平均は 85.0%であった。このうち、a 病名(110 題)の平均正答率は 86.4%、b 病名(32 題)の平均正答率は 84.0%、c 病名(5 題)の平均正答率は 74.3%である。難しい病名に関する正答率が低いことは一般論として頷ける。その一方で、a 病名に関する設問だけを選んで

分析しても、100%近い正答率から 10%台の正答率まで分布していた。本研究報告書では公表していないデータであるが、第 118 回国試医学各論 a 病名問題で、正答率 60%を切る問題が 10 数題、正答率 60~80%の問題が 20 数題あった事実からみても、a 病名であるから国家試験が易化するという主張はあたらない。設問の難易度は扱う疾患だけで決定されない。例えば、a 分類の疾患に関して確定診断をつける専門医レベルの検査法や術式等の判断など、診療参加型臨床実習で学修する水準を超えた専門医レベルとした場合などが該当する。

2. 第 118 回国試「医学各論」における回答肢に現れた a,b,c 病名の分析

- ・ 出題内容や正答肢が a 病名であっても、他の選択肢が c 病名では難易度が高くなることが想像される。このため、回答肢に病名がある設問をすべて検討し、a,b,c 病名のいずれであるか、もしくはそれ以外であるかを検討した。a,b,c 病名以外の場合、備考欄にある病名かどうか併せて検討を加えた。一例を挙げると、第 118 回の医学各論問題に乳癌の診断に関する問題がある。鑑別疾患として用意された選択肢のうち正答肢の乳癌は a 病名、4 つの誤答肢のうち乳腺症は b 病名、乳腺線維腺腫は c 病名で、葉状腫瘍と慢性乳腺炎は R6 出題基準に掲載されていなかった。
- ・ 図 2 に示したとおり、回答肢に出現した率は a 病名が 47.9%、b 病名が 23.0%、c 病名が 6.8%であった。c 病名を正答肢として出題した問題は 1 問のみで、c 病名を回答肢に用意した 20 問題の 5%だった。ここでも c 病名の使用頻度が低いことは望ましい一面と考えられる。一方で、回答肢が a,b,c のいずれの病名でもない出題が 29 題(全回答肢の 9.8%)に認められ、うち 6 題(29 題のうち 21%)は正答肢として出題されていた。回答肢が備考欄にある病名から選ばれた出題は 37 題(全回答肢の 12.5%)に認められ、うち 17 題(37 題のうち 46%)が正答肢であった。
- ・ 図 3 の左側は、a,b,c 病名に該当のない具体的

な回答肢を示している。医学生が知っておくべき疾患と考えられるものもある一方で、専門医レベルと思われる疾患も存在する。出題基準に掲載されていない(a,b,c 病名にあたらない)病名が選択肢に表現されることは本来望ましくないと考えられるが、出題者の立場からすると、5 つの病名をすべて a,b,c から選ぶことには困難が伴うと想像される。先述の乳癌の例もそのことを示しており、出題基準というルールブックだけで国試に出現する病名すべてに対応することが妥当なのかどうか議論を要するところである。

- ・ 回答肢が備考欄から提示されることが 37 題あり、全回答肢の 12.5%を占めた。そのうち正答肢としての 17 題は 37 題のうちの 46%であり、a 病名での正答肢率 23%を大きく上回っている。a,b,c,該当なし、備考欄の回答肢の 5 つのカテゴリーを単純に割り算した正答肢確率 20%を大きく超えていることは、医学各論の出題基準そのもの(特に備考欄のあり方について)を考慮すべきであることを示唆している。なお、備考欄から回答肢が使用された例(図 3)を見ると、いくつか a 疾患ではないかと思われる病名(Basedow 病や C 型肝炎など)が認められる。後述するように、医学各論の中には大・中・小項目の分類に収めないで独自に備考欄に病名を積極的に記載している領域があることが原因となっている。次回の出題基準改定の際には、統一した記載が必要である。
- ・ R6 版を含め、出題基準改定の委員を現役の国家試験作成委員が務めている構造から、備考欄になんらかの記載をしておくことがその領域の出題者にとって回答肢(しかも正答肢として)確保のための一方策となっている可能性がある。この状況を踏まえて、備考欄の調査が必要であると考えて新たな検証に移ることとした。

3. R6 出題基準に記載された医学各論の備考欄に関して

- ・ 第 116 回から第 118 回までの 3 年間の国試医学各論では、それぞれ備考欄から第 116 回で 58 問、第 117 回で 71 回、第 118 回で 71 回

の出題がされていた。R6 出題基準の改定を経ても、第118回の備考欄からの出題は減少しておらず、出題にあたって備考欄が出題における「抜け道」のようになっていることが危惧される。

- ・このような懸念から、備考欄全体の傾向を把握するために[追加病名][検査・術式・治療等][説明・付加症状]の3つのカテゴリーに類型化して検討した。その結果、図4に示すとおり、備考欄での476の記載のうち283項目(59.5%)は「追加病名」であり、出題者側が備考欄に病名を記載しておこうとする傾向が認められた。
- ・「追加病名」を多く記載する大項目領域のうち、VI消化器・腹壁・腹膜疾患領域では小項目の病名を更に細分化する傾向があった(例:④十二指腸腫瘍の備考欄に十二指腸腺腫・癌、十二指腸乳頭部腫瘍)。IX神経・運動器疾患領域においても細分化の傾向(例:③ミトコンドリア病の備考欄にKearns-Sayre 症候群、Leigh 脳症)とともに代表例の例示(例:③四肢骨折の備考欄に大腿骨近位部骨折、橈骨遠位端骨折)などが見られた。X内分泌・代謝・栄養・乳腺疾患領域)では③リピードシスの備考欄にTay-Sachs 病、Gaucher 病、Niemann-Pick 病、Fabry 病の比較的専門的な病名記載が認められた。もともと③リピードシス自体がc病名でありその下にさらに病名を追加すること自体に意味があったのか問う必要がある。
- ・一方で、先に述べたとおりV心臓・脈管疾患の[大項目7]血圧異常、[中項目B.]二次性高血圧、[小項目③]内分泌性高血圧症の備考欄に、原発性アルドステロン症、Cushing 症候群、褐色細胞腫、先端巨大症、甲状腺機能亢進症が記載されていて、この領域のみ層構造を逸脱している。そのために備考欄に病名が入ってしまう事態も見られた。遅くとも次回の出題基準改定までには他領域と揃えた層記載に変更すべきである。なお、この状況はこの領域のみに例外的なものであった。
- ・「検査・術式・治療」等の備考欄記載が多かったI先天異常、周産期異常、成長・発達異常領域では、胎児心拍数陣痛図<CTG>やApgarス

コアなどが備考欄に記載されて頻回に出題されている。これらは医学総論のⅦ,Ⅷに記載されていて医学各論の備考欄には二重記載の状態となっており、削除して問題ないと考えられる。Ⅲ皮膚・頭頸部疾では眼内レンズ挿入術や鼓室形成術など、V心臓・脈管疾患ではBlalock-Taussig 短絡手術、Rastelli 手術など専門医レベルと思われる術式が掲載されている。術式などは卒業レベルで問うことを避けることが望ましく、領域間の整合をとる必要がある。

- ・「付加説明」は、領域によって備考欄の使用理由がさまざまであり、その内容と深さに一定性が認められなかった(例:皮膚疾患 ①環状肉芽腫の備考欄に糖尿病、④反回神経麻痺の備考欄に嗄声。①肥大型心筋症の備考欄に心臓(性)突然死。③溶血性尿毒症症候群<HUS>の備考欄にVero 毒素など)。原因となり得る疾患を記載していたり、本来医学総論に掲載されるべきものであったり、メモ書き的に使用されていたりすることも散見される。
- ・これらの状況を総合して、備考欄に関して次のように推奨したい。

- 1) 本来医学各論の小項目構成は「病名」を中心とした記載するのが原則である。検査・術式・治療法・所見・説明などを記載することは控え、領域間で整合をとることが望まれる。
- 2) 備考欄に病名を記載することを停止し、備考欄の病名に関しては必要なものを小項目に編入するか、出題基準から除くこと。特にc病名の下に更なる細病名を置くことは避けるべきである。
- 3) 医学各論の出題基準における病名は、どの疾患についてどの深さで設問するかのガイドライン機能に限定する等、基準の位置付けに見直しが必要である。全ての回答肢を網羅できることを出題基準に求めるべきではない。

4. 中心的な出題委員、主要な出題委員に対するインタビューおよびアンケートに関して

[中心的な出題委員へのインタビュー]

- ・ 医師国家試験作成を中心的に行っている 4 名のインタビューでは、病名の a,b,c 分類が与えたポジティブな影響が語られた。特に、出題基準逸脱や難易度の高い問題に対して、他の委員からレベル分類を根拠としたピアレビュー的なコントロールが行い易くなったという指摘は今回の改定を支持するものである。
- ・ 中心的な出題委員からは、医学総論＋医学各論の合否水準が上昇した(易化している)現状に対する不安と言える発言があった。識別指数が低下するという試験理論によるものだけでなく、80%を合否基準とする必修とすでに合否水準が 76.7%まで上昇した医学総論/医学各論の差が見当たらないことに対する強い戸惑いが伝わった。国試が CBT 化されれば、必修問題は不要になるというコメントもあった。またインタビューでは「必修と医学総論(＋医学各論)はそれぞれ必要なのか」「多くの試験委員から必修の必要性・意義が不明という意見があった」という必修に対する考えが述べられるとともに、「試験後に合否基準が示される現行のハイステークス資格試験(医師国家試験)は妥当なのか」という疑問も呈された。彼らの経験と重い職責を考慮する時、これらの意見は深く受けとめ、尊重されなければならない。

[主要な出題委員へのアンケート]

- ・ 主要な出題委員とは、現実に医師国家試験の問題作成においてブラッシュアップに関わる数十名の委員であるが、その人数と氏名は公表されていないため、報告書ではこのような表現をとることとした。彼らの意見はウェブアンケートで集計されたが、一人残らず全員から回答を得られたことは試験委員の意識の高さを示すもので敬意を表したい。悉皆性のある調査となり信頼性も保証されることとなった。
- ・ アンケートの結果は章末に記載したので参照されたい。ここでは主な結果(図 6,7,8)をもとに考察を行う。

- ・ 主要な出題委員にとっても、改善検討部会報告書を意識していたと答えた者は 54.8%(ある程度意識を含めても 85.8%)に留まった(図 6)。一方でレベル分類への意識は(ある程度を含めて)100%だったことから、作問に関しては「出題基準における記載内容」がより具体的な意識づけにつながっていると考えられる。出題基準への正しい記載と運用が重要である事実が改めて認識された。前述のとおり出題基準改定作業の実務は、従来現役の出題委員に委ねられる構造であり、出題基準が肥大化する方向性が危惧され、実際備考欄などでその証左が認められた。出題委員経験者から医師国家試験の全体像を掴んでいる(例えば今回調査における「中心的な出題委員」等)人材に改定作業の統括を任せるなどの方法を探ることも一法と考える。
- ・ 主要な出題委員に対する「必修と医学総論の重複」について、半数近い 48.9%が統合・整理など検討が必要と回答した(図 7)。「CBT 化により必修問題を作成する必要性がなくなるのではないか」という内容のコメントが複数あり、出題委員のなかにも国試の CBT 化の方向性を予想している者が存在している。
- ・ 必修の位置付けでは、特に問題ないが 77.8%であり一見必修に問題がないように見えるが、見直しが必要とした 22.2%に多数のコメント記載が認められている。これらを再掲すると「易しすぎて問題の体をなさないこともあり、医学総論との共存の検討を」「ほぼ全員が正答する試験を作る意味合いを問う」「一般教養の事項 A は削除を」「(臨床実習前共用試験)CBT 開始から年月も経ち、必修の必要性について議論が必要 → 総論と各論のみとした方が分かりやすい」「高い正答率を求められ道徳的に回答できる極端な易問が多く出題されている」などがあった。出題委員のカテゴリーや内容(非公表)をこの報告書で触れることは避けるが、専門分野に特化した医学各論中心の委員と全般的な共通分野を受け持つ必修・医学総論等の委員とに温度差があることが原因のひとつと推察される。問題なしとする率よりも問題があるとする具

体的なコメントに重きをおいて必修のあり方を検討する必要がある。

- ・ 医学総論の位置付けにおいても、特に問題ないが 77.8%で同様の傾向がみられた。ここでも、医学総論の見直しが必要とする意見には「必修との差が分かりにくい」「公衆衛生的な内容が多すぎるように感じる」「各論との違いが難しく、タキソノミーⅡ,Ⅲでは治療等となり診療科をまたぐより各論疾患に対しての対応を聞きがちになる」などがあった。必修へのコメントと合わせ、両者の課題が認識される。
- ・ 医学総論における保健医療/公衆衛生領域の「問題数」について 62.2%の委員から見直しが必要とされた(図 8)。この背景として、医師国家試験が平成 30 年に施行された第 112 回以降 500 問から 400 問に問題数が削減された際に、保健医療/公衆衛生系の問題が呼応して削減されなかったことが一因となっている。現状の医学総論 150 問のうち I 保健医療論が約 13%、Ⅱ予防と健康管理・増進が約 17%と比較率が高いまま残っている一方で、共用試験ガイドブックには「A 医師として求められる基本的な資質・能力」と「B 社会と医学・医療」の二項目を合同してなお 10%程度の出題を行うとしか規定されていない。この領域の設問には、臨床実習前に問うべきものもあれば臨床実習を経験してから問うべきものもある。臨床実習前の共用試験と医師国家試験の整合をとってこなかったシステム上の課題がアンケート結果に結びついていることに留意しなくてはならない。実習前後の学修者評価がどうあるべきかという制度設計を含めた俯瞰的な戦略が必要と考えられる典型例と言える。

5. 研究の全体像に関して

- ・ この研究は、医師国家試験出題基準における医学各論のあり方について、主に第 118 回医師国家試験の情報をもとに検証を行ったものである。検証の目的は、一義的に近々開催が想定される次期医師国家試験改善検討部会への報告である。
- ・ 他方、医学各論を含む医師国家試験出題基準

は、医学生 of 学修者評価に関わるもので、我が国の医療を中心となって支える医師の育成・評価という広い視野のもとに考察を加える必要がある。その視点から、少なくとも医学生が身につけるべき資質・能力を記載した令和 4 年版医学教育モデル・コア・カリキュラムの内容との整合は必須となる。特に医学各論では病名を基本とした記載となっていることから、コアカリに記載された病名との整合は必須である。

- ・ 臨床実習前の知識の総括的評価となる共用試験 CBT は、公的化された現状を踏まえ、学修者評価において医師国家試験と総合的に作問や継続的改良を行うシステムを構築することが望まれる。そのことは、医療を受ける国民と学修の主体である医学生にとって、よりよい成果をもたらす。

6. 研究の limitation

- ・ この研究は、医師国家試験作成の運営にあたっていた経験をもつ研究者 2 名が中心となって行われた。現行の試験作成とは独立して利益相反はないものの、医師国家試験の内部事情に詳しい背景を持つことから、研究者の意見に基づくバイアスが生じる可能性がある。
- ・ 卒前後における一貫した学修者評価の面から、医師国家試験での医学各論という部分に関する研究成果は医師育成における全体像を見極められていない可能性がある。考察と結論に述べたように、医学各論と医学教育モデル・コア・カリキュラム(特にその病名等)および臨床実習前の共用試験(CBT)などとの整合が課題として残されている。

【今後への提言】

- 医学各論の小項目は「病名」を中心とした記載に統一することを原則とし、備考欄の整備を含めて領域間で整合をとる必要がある。
- 医学各論における出題基準は、どの疾患についてどの深さで設問するかガイドライン機能に限定する等、出題基準の位置付け

に関して見直しを行うことが望まれる。

- 医学各論は領域別病名を基盤とした記載となっていることを踏まえ、医学各論における病名のさらなる整備とともに、医学生が身につけるべき資質・能力を記載した医学教育モデル・コア・カリキュラム(特にその病名等)との整合を図ることが重要である。

E. 結論(第 118 回国試験検証)

1. 令和 6 年版医師国家試験出題基準における医学各論の病名の a,b,c 分類は概ね適切に機能したと考えられる。

- ・ 第 118 回医師国家試験において a 病名からの出題が 8.6 ポイント増加(64.7→73.3%)した。
- ・ 医師国家試験の委員インタビュー/アンケートからも a,b,c 分類は有効に機能していることが確認された。
- ・ 出題病名だけでなく回答肢における a,b,c 病名などについて確認したところ、a,b,c 病名に該当しない 29 病名(全回答肢病名の 9.8%)や備考欄にある 37 病名(同 12.5%)がある等の課題が認められた。

2. 備考欄に a,b,c 分類の適用を受けない病名が存在している。

- ・ 476 の備考欄記載のうち 283 項目(59.5%)は「追加病名」であり、62 項目は「検査・術式・治療等」、131 項目が「説明・付加症状」に大別された。
- ・ 備考欄からの出題の殆どは「追加病名」が占め、第 116 回～第 118 回国試で 58,71,71 題が出題された。

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

