

試験問題の作成の手順について

専門医制度委員会委員長

今井 裕一

はじめに

現在、専門医制度の大改革が始まっている。基本領域学会（日本内科学会、日本小児科学会、日本外科学会、日本泌尿器科学会）では、2017年度からプログラム研修が開始される。サブスペシャリティ学会である日本腎臓学会はその3年後からスタートする予定になっている。大きな変更点は、専門医の認定に関しては、日本専門医機構が行うことになる。また、専門医の更新に関しても新たな基準が求められている。

これまで専門医制度委員会への協力として、試験問題作成に関して2単位、最大4単位が付与されている。これらの問題を基にして作成されたセルフトレーニング問題に合格すると5単位が付与されている。質の高い試験問題を作成することは、専門医および指導医のレベルを維持することに不可欠である。

しかし、これまでどのような手順で問題を作成するのか、明確でなかったので今回文章にまとめた。参考になれば幸いである。

形成的評価と総括的評価

教育現場において最も重要な点は、教わる側をどのように評価するかということである。正当な評価がなければ、次の改善策、レベルアップも生まれない。逆に言えば、評価をすることが、成長のための第一歩なのである。教える側にとっても教わる側にとってもこの認識を共有する事が大切である。

評価には2つの方法があり、一つは、いろいろな基準項目に従って評価し、それを相手にフィードバックすることで教わる側の成長あるいは改善を促す方法である。これを「形成的評価」と呼んでいる。特に実際の臨床現場で若い医師を教育する場合に使用される。もう一つは、総括的評価である。これは、研修あるいは学習プログラムの終了時に到達目標をどの程度達成できているかを評価するものである。最終的な評価になるが、学習者個人を評価することはもちろんであるが、受験者全員の評価を統計的に処理することによって、教育計画、実施方法（試験問題の質と量）の妥当性を再評価し、システムを改善することができるようになる。しかし適切な試験問題でないと、総括的評価を行う意味は消滅してしまう。総括的評価は、大学教育、卒後教育、専門医教育にとって不可欠な評価法である。今回は、総括的評価として重要な試験問題の作成方法について概説する。

総括的評価における試験問題の作成方法

1. どのような能力を評価しているのか？

設問のタイプによって、「どのような能力」を評価するのかが決まっている。臨床能力は、「単なる暗記（記憶力）」「問題点の解釈能力」「問題点の解決能力」に大別されている。専門医として必要な能力は、「問題点の解決能力」であり、それが到達目標に達しているかを評価することになる。上記のそれぞれの能力をタキソノミー（Taxonomy）という医学教育用語で表現している。

2. 試験問題のタキソノミー (Taxonomy)

1) タキソノミー 1 recall 単純記憶型

設問に関連する知識の有無を単純に問う問題であり、知っていれば正解することができるので、このような問題が多いと丸暗記すれば合格することになる。このような学習方法も確かに必要ではあるが、暗記中心の試験問題となり受験生の問題解決能力を評価することはできない。

.....

例 1. X 染色体劣性遺伝の疾患はどれか。2 つ選べ。(正解：a, c)

- a. Fabry 病
- b. 多発性嚢胞腎
- c. Alport 症候群
- d. Liddle 症候群
- e. Gitelman 症候群

.....

【注釈】「1 つ選べ」と「2 つ選べ」では、明らかに後者が難しくなる。全体の出題の難易度を考えて、どちらにするかを最終調整する。専門医試験レベルでは「2 つ選べ」は 15～20 % が妥当と思われる。

2) タキソノミー 2 interpretation 問題解釈型

設問の意味を理解して回答を選ぶ問題である。多くは、疾患を想起する能力の有無を問うものである。さらには、その疾患に関する問題点を質問することになる。

.....

例 2 【問題文】

25 歳の女性。2 週間前に扁桃炎に罹患した。昨日から顔面と下肢に浮腫を自覚して受診した。尿検査を行なったところ以下の結果を得た。

pH 6.2, 蛋白 2+, 潜血反応 3+, 尿糖 陰性, ケトン体 陰性,
尿沈渣：赤血球 40/弱拡大, 白血球 30/弱拡大, 変形赤血球多数, 赤血球円柱多数

【設問】検査で異常となるものはどれか。

- a. HbA1c 上昇
- b. 抗 DNA 抗体
- c. 血清 IgA 上昇
- d. 血清補体低下
- e. 尿中 Bence Jones 蛋白

.....

【注釈】問題文で疾患を正しく想起できなければ、正答確率は 20 % 程度である。

上記の問題文では、キーワードとして、「2 週間前の扁桃炎」、「糸球体腎炎性の尿所見」をピックアップして、急性糸球体腎炎（溶連菌感染後糸球体腎炎）を想起させる。

その際の検査値異常を予測することになる。設問としては、予想される「身体所見」、「検査」、「診断名」などがある。この場の問題文は、実際に受け持った患者のデータではなく、典型的な症例のデータに変

更・調整する。

- a. HbA1c 上昇←糖尿病？
- b. 抗 DNA 抗体←全身性エリテマトーデス？
- c. 血清 IgA 上昇←IgA 腎症？
- d. 血清補体低下←急性糸球体腎炎？
- e. 尿中 Bence Jones 蛋白←腎アミロイドーシス？

3) タキシノミー 3 resolve 問題解決型

単に問題文の解釈だけではなく、選択肢の意味も理解して、回答を探す必要がある問題である。臨床現場に則した問題である。実際に経験した症例、あるいはカンファレンスで討論した症例を想起し、妥当な対応を問う問題である。

例 3【問題文】

48 歳の女性。発語障害，失見当識および左上肢麻痺があり救急外来を受診した。

身体所見：体温 37.6°C。貧血あり，上肢と腹部とに紫斑を認める。

頭部 CT 検査では，明らかな出血は認めなかった。外来で検査結果を待っている間に，症状が軽快し，意識清明，発語障害，左上肢麻痺も消失した。

尿所見：蛋白 (+)，潜血 (+)。

血液検査：Hb 6.8 g/dL，白血球 8,490/ μ L，血小板 5.0 万/ μ L，

プロトロンビン時間 103 %，活性化部分トロンボプラスチン時間 31.1 秒，

血漿フィブリノーゲン 301 mg/dL，FDP 30 μ g/mL (基準 \leq 10)。

BUN 40 mg/dL，Cr 1.8 mg/dL，尿酸 7.5 mg/dL，AST 51 U/L，ALT 32 U/L，LDH 1,408 IU/L，

【設問】適切な処置はどれか。2 つ選べ。

- a. 血漿輸注
- b. 血漿交換
- c. ヘパリン投与
- d. アスピリン投与
- e. 濃厚血小板<PC>輸血

【注釈】問題文を読んで，キーワードとして「一過性脳虚血症状」，「尿異常：糸球体病変？」，「貧血，LDH 上昇→溶血性貧血？」，「線溶系亢進」「腎機能低下」がピックアップされる。もっとも可能性の高い疾患として，「溶血性尿毒症症候群／血栓性血小板減少性紫斑病」が想起される。類似した疾患として，1) 播種性血管内凝固症候群，2) 血小板減少性紫斑病などがあるが，それらをダミーとして選択肢に入れる。臨床経過と検査値から疾患を推測し，それに対しての適切な治療法を選択させることで，受験者の問題解決能力を評価する。

-
- a. 血漿輸注→ADAMTS 13 活性が低下している場合は、有効である
 - b. 血漿交換→TTP では有効である。
 - c. ヘパリン投与→DIC ではないので無効である。
 - d. アスピリン投与→無効である。
 - e. 濃厚血小板<PC>輸血→TTP では禁忌になる。
-

3. 問題の種類

1) 一般問題

一般問題とは、専門医として必要な知識（解剖、生理、薬理、症状、検査データ、手術手技、病態）を単純に問う問題で、タキノミー1（単純記憶型）に相当している。まれな疾患で経験はしていなくても、知識があるかどうかを評価する際にも使用する。

2) 一般問題の作成手順

- ①キーワードを想定する。研修カリキュラムに含まれている用語を選択する。

この点は決定する前に確認しておく。

（例：尿細管機能、ホルモン）

- ②設問と正解数を決定する。

（例：尿細管に直接作用するホルモンはどれか。2つ選べ（b, d））

- ③選択肢を単語の短い順、あるいはアイウエオ順などの順に並べる。

この際の選択肢は、すべてがもっともらしいものを選ぶ。明らかにカテゴリーの異なるものは、ナンセンス選択肢であり、そのことだけで正解の確率を上げてしまう。

（均等な場合：20% →1 選択肢が明らかに除外される場合：25% →2 選択肢が明らかに除外される場合：33% →3 選択肢が明らかに除外される場合：50%）

- ④選択肢はできるだけ短く単語1つにする。

長い文章になると、論理が入り無用な混乱を生じる。

例4 尿細管に直接作用するホルモンはどれか。2つ選べ（正解：c, d）

- a. レニン
 - b. プロラクチン
 - c. アルドステロン
 - d. バゾプレッシン
 - e. エリスロポエチン
-

例5 MPO-ANCAが陽性の場合、可能性の高い疾患はどれか。2つ選べ。(正解:c, d)

- a. 混合性結合組織病
- b. Wegener肉芽腫症
- c. 顕微鏡的多発血管炎
- d. Churg-Strauss症候群
- e. 全身性エリテマトーデス

2) 臨床問題

臨床問題とは、年齢、性別を記載し実際の臨床カンファレンス形式で症例を提示し、設問に答えるものである。問題解決能力を重視する問題、解釈あるいは問題解決を求める問題であることが望ましい。今後の専門医試験に関しては、タキソノミー1の出題数を減らし、タキソノミー2、あるいは3を増加させる必要がある。

問題文は、以下のように記載するルールになっている。

〇歳の男性・女性。〇〇があり、受診した。

現病歴：5年前、健診で高血圧を指摘されたが放置していた。2カ月前から頭痛が出現し徐々に増悪してきた。数日前から右上肢に力が入りにくく感じるようになった。

身体所見：意識清明。身長〇〇cm、体重〇kg。脈拍〇/分、整。血圧〇〇/〇〇mmHg。眼瞼結膜 貧血あり、なし、眼球結膜黄染 あり、なし。胸腹部異常なし。四肢に浮腫を認めない。神経学的所見異常なし。

尿検査：蛋白3+、糖(-)、潜血1+。尿沈渣：

血液検査：赤血球〇〇〇万、Hb〇g/dL、Ht〇%、白血球〇、血小板〇万。

生化学検査：血糖〇mg/dL、総蛋白〇g/dL、アルブミン〇g/dL、尿素窒素〇mg/dL、クレアチニン〇mg/dL、AST〇IU/L、ALT〇IU/L、LDH〇IU/L(基準176~353) Na〇mEq/L、K〇mEq/L、Cl〇mEq/L、さらに、必要な検査結果を追加する。

3) 連問の作り方

連問とは、問題文は同一であるが、設問が複数ある場合を指している。設問は、①臨床症候を問う設問、②検査成績の解釈を問う設問、③起こりうる事態を予測する設問、④治療法を問う設問など、それぞれが独立している必要がある。特に前の問題が解けないと後の問題が解けないようになっていると悪問である。

2. 連問作成の手順

臨床問題の問題文を作成し、独立した設問を決定する。(例：①検査の異常、②治療法：食事療法)

例 6

【問題文】(例 2 に検査成績, 腎生検病理組織を追加する。)

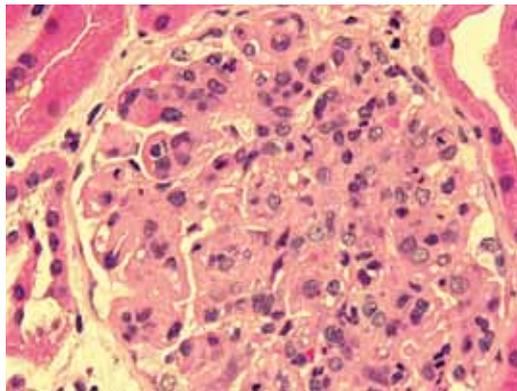
25 歳の女性。2 週間前に扁桃炎に罹患した。昨日から顔面と下肢に浮腫を自覚して受診した。尿検査を行なったところ以下の結果を得た。

pH 6.2, 蛋白 2+, 潜血反応 3+, 尿糖 陰性, ケトン体 陰性,

尿沈渣: 赤血球 40/弱拡大, 白血球 30/弱拡大, 変形赤血球多数, 赤血球円柱多数

生化学検査: 総蛋白 7.2 g/dL, アルブミン 4.1 g/dL, BUN 28 mg/dL, 血清クレアチニン 1.2 mg/dL, Na 136 mEq/L, K 4.6 mEq/L, Cl 100 mEq/L。

入院 3 日目に施行した腎生検の糸球体像を示す (図 HE)。



【設問 1】検査で異常となるものはどれか。1 つ選べ。(正解: d)

- a. HbA1c 上昇
- b. 抗 DNA 抗体
- c. 血清 IgA 上昇
- d. 血清補体低下
- e. 尿中 Bence Jones 蛋白

【設問 2】食事療法で正しいのはどれか。1 つ選べ。(正解: c)

- a. 総エネルギー 25 kcal/kg (標準体重)/日
- b. 蛋白 0.5 g/kg (標準体重)/日
- c. 塩分 7g/日
- d. 水分制限せず
- e. カリウム制限

【注釈】問題文の臨床経過と簡単な検査結果から、「急性糸球体腎炎」を想起し、病理組織標本で「管内増殖性腎炎」を確認できるかという能力を評価する。さらにこれに基づいて治療法を選択させる問題である。このような臨床現場に則した問題が好ましい。

試験問題の評価

1) 設問の正答率

通常は、正答率が60～70%の問題が、難易度が普通であると判断されている。正答率80%以上の問題は、必修問題レベルのやさしい問題である。一方、正答率50%未満のものは、難しい問題である。難、普通、易の問題数が3等分されていれば、平均的には、受験者全体では普通レベルの試験問題になる。最終的には、試験委員会で、難、普通、易の比率を微調整することになる。特に、臨床研修を真摯に行い、十分な実力をもった受験生は、高い得点となり、臨床経験が不十分で暗記のみの受験生の得点が低くなるように問題を調整する。正答率が、20%未満の問題は、削除問題の対象となるが、最終的に採用するかどうかは、試験委員会の判断になる。

2) 設問の識別指数

受験者の総合得点の高いグループと低いグループの正答率に基づいて、1つ1つの設問で識別指数を算出する。この数字が、おおよそ0.2以上であれば良問であると判断できる。0.4以上であれば非常に優れた問題である。ただし、全体の正答率が90%程度あるいは20%未満になると、識別指数が低下することは明らかである。識別指数がマイナスになるような問題は、削除問題に相当するかどうか試験委員会で審議される。

以上、試験問題の作成の手順をまとめたので、是非参考にして、作成した問題を専門医制度委員会（事務局）に送ってください。

チェックリスト

出題基準・レベル等について

- 出題基準（研修カリキュラム）に含まれているか。
- 出題レベル（難易度）は適切か。
- 単に知識の想起により答えられる問題ではなく、問題解決能力を重視する問題となっているか。

出題内容について

- 臨床研修の成果が反映される問題（臨床研修でしばしば経験する症候・病態・疾患）となっているか。
- 日常遭遇することがないようなまれな疾患ではないか。→この場合は、知識のみを問う
- 数値に関する設問では、診療するにあたり、常に記憶していなければならないものか。
- 成因や機序を問う設問の場合、意見の分かれているものではないか。
- 性差、年齢差に関する設問の場合、差のあることが広く知られているか。
- 人名を冠した疾患、症候群、検査法、手術術式を問う問題では、必ず知らなければならないものか。

記載の仕方

- 人名は原語で記載されているか。
- 細菌の学名はイタリック体で、ウイルスはローマン体で表記されているか。
- ガイドラインを参照させる問題では作成年も記載されているか。
- 図はオリジナルであるか。著作権を侵害している可能性はないか。
- 一般問題では専門的であっても必須の知識や判断力を問う問題となっているか。
- 受験者の60%以上が答えることができるような問題となっているか。
- 臨床問題では、年齢、性別を記載して実施形式で症例を示し、設問に答える問題となっているか。

選択肢について

- 正解肢として可能性を否定できない選択肢はないか。
- 同一問題内の選択肢はすべて同じカテゴリーのものであるか。
- 「必ず」とか、「ありえる」などの言葉が入っていないか（選択肢で、必ずという用語は100%誤りであり、「——ありえる」は100%正しいので、使用しない）。

連問について

- 特に同一連問の中の各問題はできる限り独立した内容のものとし、前の問題が解けないと後の問題が解けない様な問題になってはいないか。