

# 座談会 試験・評価を学習者の味方に!!



「医学教育モデル・コア・カリキュラム」(以下、コアカリ) 令和4年度改訂版では、「学修評価」についての節が独立して設けられ、医師として求められる資質・能力を適切に評価するための考え方、評価方法等、幅広い内容が扱われている。

本紙では自治医科大学で医学教育研究に従事する松山氏を司会に、コアカリ令和4年度改訂版の「学修評価」の節を取りまとめた伊藤氏、「医学教育モデル・コア・カリキュラム等の次期改訂に向けた調査・研究 医学チーム」で副座長を務めた錦織氏、教育学の専門家である齋藤氏を迎えた座談会を企画。今後の医学教育における学習者評価の在り方を考えた。

**松山** 本日はお集まりいただきありがとうございます。コアカリ令和4年度改訂版では学修評価という節が新しく設けられました。医学教育における学習者評価の重要性が認識されつつある今、コアカリ作成に携わったメンバーも交えて議論ができることを非常にうれしく思っています。

**齋藤** コアカリ令和4年度改訂版を拝読しましたが、医学教育において何のような基準で評価しなければいけないのかとの視点が明文化され、教育現場の判断に役立つ情報がわかりやすくまとめられていると感じました。

**伊藤** ありがとうございます。私が学修評価の節の取りまとめを担当しました。コアカリの読者は医学教育を専門にされていない方も多いため、評価ツールなどの各論よりも、総論的な理解を深めてもらうことをめざしました。

例えば、基本的な医学教育における評価の概念として紹介したのが“Millerのピラミッド”(MEMO)です。医学教育で実践的な能力の評価をするに当たっては、模擬的な環境も含めた場で自らの能力を行動で示す力である Shows how、実際の診療現場において実践する能力である Does に対する評価が主体となります。

**松山** コアカリでは資質・能力の評価についての基礎も説明されていましたね。医師として求められる資質・能力は複数の多面的な能力から成るものであり、その能力は知識や技能、価値観、態度などの要素を含む観察・評価可能

な能力(コンピテンシー)から構成されます。1つの評価方法で学習者のコンピテンシーを完璧に評価することは不可能です。筆記試験、OSCE等の実技試験、診療現場での観察評価(WBA: workplace-based assessment)、ポートフォリオ等で多面的に、妥当性を考慮しながら評価される必要があります。

**錦織** 学習者評価は各大学の特色などさまざまな影響を受けるものであり、決まりきった正解はありません。そのためコアカリの中では、資質・能力の評価について考えてもらいたい問いをいくつか読者に投げかけています。各教育機関で問いの答えは考えていただきたいと思っています。今後の問いからコアカリがどう発展していくのか、今から非常に楽しみです。

## 学習者にとって評価は敵なのか

**松山** 学習者評価の議論に際して、まずは日本の医療教育における「評価」のイメージを考えたいです。私は、医学に限らず日本の教育全般で「評価=試験」という先入観があるのではないかと感じています。集団で一斉に行い、その成否が学習者のその後の人生を左右する、医師国家試験のような試験が評価のトップに位置付けられている印象を抱きます。先生方のイメージはいかがでしょうか。

**伊藤** 全く同意見です。試験は試験でも授業内で行う小テストのような軽い

ものではなく、入学試験、卒業試験のような格式ばったものがイメージされているように思います。

**齋藤** 合否判定や選抜の場面でのみ評価が意識されている、という印象でしょうか。確かに、学習過程で行われ、学習者が目標を達成するために不足している点を気づかせ、改善を促すための「形成的評価」より、学習過程の終了時期に行われ、学習者が目標に達しているか否かの判定に用いられる「総括的評価」が圧倒的に重視されている雰囲気はあります。

**錦織** 国試はまさに医師になる人を選抜する総括的評価の場ですよね。

**伊藤** そうした文脈での試験は「傾向と対策」といった言葉とセットで語られます。立ち向かって越えるべき、相対するものであって、自分の味方になるものではないと考える方が多いと思います。

**錦織** 試験を課す先生はロールプレイングゲームの敵キャラクターのようなイメージですよ。だから楽々倒せるような問題を出してくれる先生が良い先生と思われていたり(笑)。

**齋藤** 良いとえですね。医学教育現場では、1つでも単位を落としたり留年する可能性があったり、国試と同じ形式で持ち込み禁止のペーパー試験を行ったり、インパクトの大きい国試のための厳しい対策が日ごろから強く意識されているように見受けられます。「敵を倒すために頑張れ」と1年生のころから常に学習者に発破がかけられている印象です。

**松山** 私が懸念しているのは、国試という厳格な総括的評価とその対策のための学習が、臨床能力が高い医学生を育て、評価するという本来の目的からかけ離れたものになってはいないかという点です。ペーパーテストに強い学生を作り選抜するという、敵を倒す点のみに注力した学習過程になっては本末転倒ではないでしょうか。

評価は敵ではなく、学習者の成長を支援する「味方」です。「評価を味方に」という視点で、以降の議論を進めていければと思います。

## 形成的・総括的評価を兼ねる Workplace-based assessment

**齋藤** 教育評価論における教育評価や学習評価の目的は、学習者を序列化したり選別したりすることではなく、学習者の資質・能力の発達を保証することと、教育活動を改善していくことです。評価によって学習者に自分の現在の状況を把握してもらい、成長に向けた道しるべを示して改善につなげてもらうわけです。

**伊藤** 指導・教育の中でのフィードバックに代表される形成的評価の視点です。総括的評価が目標を達成できたか否かを評価するのに対し、形成的評価は目標の達成に向けて支援する役割を果たします。まさに学習者の「味方」と言える評価だと思います。

**松山** 形成的評価としての側面も期待されている医学教育における学習者評価の概念に、Workplace-based assessmentがあります。WBAの一種である mini-CEX(註1)や DOPS(註2)などは最近評価界隈のパスワードになっています。WBAは「診療現場における学習者評価」と訳され、学習者の日常的な行動・態度を指導者が観察して評価し、評価の根拠を記録、フィードバックするものです。異なる評価者であっても一定の観点と尺度で評価できるよう、ルーブリック(註3)の作成がコアカリ令和4年度改訂版では推奨されています。

**錦織** WBAでは、診療現場での実践能力を直接観察し、Does段階における評価とフィードバックを行うため、ダイレクトに診療の質に影響すると感じています。直接的なこともあって診療の質が上がるスピードも速い。医療の質向上に大いに貢献するという意味では、医学教育の王道と言える評価方法かもしれません。

**齋藤** 指導者側もルーブリックで示されている評価観点によってどういった点を意識して教育すべきかとの軸がわかり、指導が行いやすくなります。学習者だけでなく指導者にとっても味方になる評価ツールです。

**伊藤** WBAは評価をドキュメントの形で残すのが最も労力のかかる点であ

### MEMO Millerのピラミッド

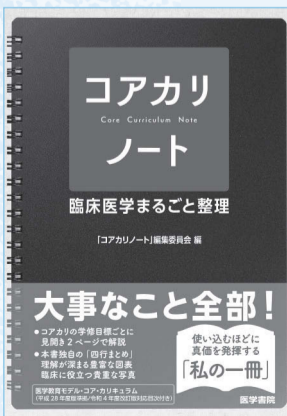
1990年に提唱された医学教育評価に関する概念図であり、下層から順に Knows, Knows how, Shows how, Doesの4層から成るピラミッドで示される<sup>1)</sup>。最も基盤にある Knows は専門職としての能力を発揮するために必要な知識、Knows how は収集した情報を分析・解釈して診療に応用する能力、Shows how は模擬的な環境も含めた場で自らの能力を行動で示す能力、Does は診療の現場で実践する能力を示す。Knows や Knows how は筆記試験、Shows how は OSCE、Does は WBA で評価されることが多い。診察などの実践的な能力の評価においては Shows how や Does の能力評価を意識する必要がある。

## コアカリノート 臨床医学まるごと整理

編「コアカリノート」編集委員会

A5 頁2,016 定価: 19,800円  
(本体18,000円+税10%)  
[ISBN 978-4-260-04864-4]

詳細はこちら



## 自分流の使い方ができる“臨床医学全一冊、

- 臨床医学全分野のエッセンスを一冊に凝縮!
- 医学生が卒業までに修得すべき事項を示した「医学教育モデル・コア・カリキュラム」に準拠し、学修目標ごとに見開き2ページで完結する構成。
- 要点を簡潔書きのスタイルでまとめた文章(左ページ)と関連図表(右ページ)が、効率的な知識の整理、記憶の定着を促します。



- 人体各器官の正常構造と機能、病態、診断、治療
- 全身に及ぶ生理的变化、病態、診断、治療
- 診療の基本 (本文執筆 安達洋祐)

医学書院

り、導入へのハードルになります。あえてドキュメントとして残す利点はどこにあると思いますか。

錦織 総括的評価への利用ではないでしょうか。指導者の頭の中だけでなされたり、口頭でのみ伝えられたりする評価は、妥当な評価が多いと思われるものの後から検証することができないため、合否判定としては活用できません。ルーブリックを用いて一定の枠組みで評価した上で文章に落とし込み、第三者からも客観的に判断できる形にすることで初めて合否判定に用いることができるようになります。

松山 同意見です。学生側から評価に対する異議が出た際にも、判断の根拠が文書として残っていないと対応できません。また記録として残すことで、その評価を学習者と指導者で共に振り返って議論する機会を作ることでも、成長につながられます。

伊藤 なるほど。また、WBAをより良く機能させるには、導入の際に何をドキュメント化するかをきちんと決めておくのと良いのではないのでしょうか。できなかったことだけでなくできたことも記録する、など運用ルールを決めることを勧めます。

齋藤 心肺蘇生や死亡宣告など、臨床で遭遇する機会が少なくWBAを活用した評価がしづらいシチュエーションもあると思います。こうした場合はOSCEのような模擬的な状況を活用した形成的評価を行うと良いかもしれませんね。

錦織 シミュレーション教育とOSCEを導入し、実技ベースで教育・評価を行う場を設けるのは力を付けるのかなり有益だと思います。臨床実習のローテート中に、各診療科でもっとシミュレーション教育やOSCEを実施できたら良いかもしれません。

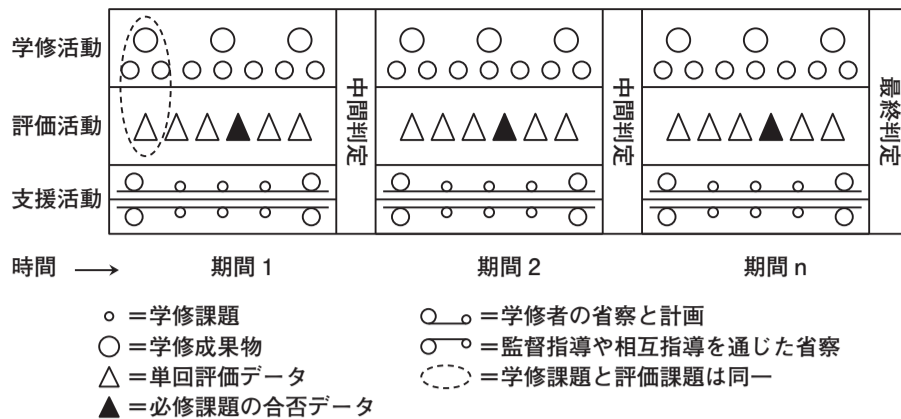
### 注目の概念

#### Programmatic Assessment

松山 医学教育において近年注目が集まっている評価概念として、2000年代前半から提唱されているPA (Programmatic Assessment)があります(図)<sup>2)</sup>。PAは、能力を経時的かつ体系的に評価し、得られた全ての結果を学びにつなげながら、合理的に十分な能力の獲得を判定すべき、という観点に基づいて考案されました。成長の過程に多様性がある学習者の能力をある特定の時点を切り取って散逸的に評価するのは妥当ではないとの考えからです。

世界の医学教育における最も大きな学会の1つであるOttawa Conference 2024に参加した際、世界的にPAの実装の潮流が来ていると肌で感じました。実装に関してなかなか最適解が見えない状況であり、PAの導入成果に関しても今後さらなる検証が必要ではあるものの、実装に取り組む大学は増えてきています。

齋藤 PAは、学習者の能力を体系的に評価し、その結果を学習につなげる



●図 Programmatic Assessment のモデル (文献2より一部改変)

学習者の成績や進捗を、ある特定の時点を切り取って散逸的に評価するのは妥当ではなく、体系的に測定し、得られた全ての結果を学びにつなげるべきであるという観点で確立された概念。学修(習)活動、評価活動、支援活動が一体化された体系を取り、次の4つの特徴を持つ。①カリキュラム内の異なる時点で複数回、多様な方法で評価を行う設計をめざす。②単回の評価結果は学習者の自己分析や形成的評価を含んだメンタリングに用い、記録として残す。③総括的評価は複数回の多様な評価結果の記録を集約して行う。④重要な判定はメンタリングや教育指導に直接かかわらない複数のステークホルダーによる合議によって行う。

新しい評価のパラダイムです。PAでは合否判定のようなその後の人生への影響が大きい評価が単一の総括的評価に基づいて行われることはなく、複数の評価方法を戦略的に組み合わせ、それらの評価情報を総体的に検討する形で実施されます。伝統的な形成的・総括的評価の区別をPAでは異なる視点でとらえます。伝統的な形成的評価も戦略的な手続きをとって行えば合否判定のような総括的評価につなげることができ、総括的評価もフィードバックでコメントを添えるなどすれば形成的に機能します。このように、あらゆる評価を連続体としてとらえてプログラム化することが特徴です。

錦織 日本の医学教育でPAを背景とする評価を実装するとすると、どういった方針が考えられるのでしょうか。

伊藤 MillerのピラミッドにおけるKnows, Knows howの段階を学んでいる低学年よりは、Doesの段階を学ぶ高学年にPAはフィットすると思います。ある程度学年が高くなってから導入したほうが良いかもしれません。

松山 その可能性はあります。また、欧米ではPAが講座や部門の垣根を越えた水平・垂直統合的なカリキュラムの中で、教育機関全体として単位認定し、進級・卒業判定する形で実施されているのに対し、日本は講座それぞれで単位認定して、それを積み上げて進級・卒業判定に用います。日本での実装には構造的な難しさが伴うでしょう。

齋藤 規則との兼ね合いはあると思いますが、各科目の成績を付けることと進級の判断とをイコールにせず、各科目での日々の形成的評価や中間テスト、期末テストの結果を資料として集積して、それを用いて科目の担当ではない教員が改めて総合的に進級判断をするという形でPAを取り入れられるかもしれません。

松山 臨床実習で、全診療科のローテートを1つのユニットとして、1年分の単位をまるごと与える、といった取り組みは日本の一部の大学でもできているので、進級を大きくくりで判

断できる可能性はあります。実装のためには、総合的に評価することのアドバンテージをいかに医学教育機関の内の指導者に提示できるかが鍵になりそうです。今まで行っていたおのおの講座での単位判定はそのまま残し、+αでPAを用いた進級判断を取り入れるイメージで導入を試みるとスムーズに進むかもしれませんね。

#### 将来の糧になる学びを保証する評価を

齋藤 評価が敵と思われるがちなのは、実際に現場で必要とされる能力を評価されている気がしないからなのかもしれません。こうした評価は倒したら終わりであり、その後生きる力として残ります。成長のための味方として評価をとらえられるように、われわれも評価を設計できると良いですね。

伊藤 中等教育までは、将来の進路が異なるためジェネリックスキルなどの教育が中心です。高等教育に来て初めて、専門職として将来現場で行うパフォーマンスに必要な能力の養成に特化した教育・評価ができるようになります。高等教育の設計者であるわれわれは、この点を意識して臨床能力を養う教育と評価を行うべきでしょう。

錦織 医学教育における評価は、日本国民が安心して医療を受けられるための質保証として必須のものです。質や均一さを過度に求めるのは問題ですが、国民に支持される医療人を育成するためにも重要なことだと思います。実現可能性に配慮しつつも、妥当性の高い評価設計をする意識は常に持つておきたいです。

松山 医学生は段階を経て学んでいくものです。医療者に必要な資質・能力を段階ごとに確実に得るための学びを保証する、そのための評価を含んだカリキュラム設計をしていくのが、われわれ指導者にとっては必要なことだと思います。本日はありがとうございます。

(了)

#### <出席者>

##### ●まつやま・やすし氏

2001年自治医大卒。15年蘭マーストリヒト大医学教育学修士課程修了、20年同博士課程修了。博士課程では学修評価の世界的権威であるCees van der Vleuten教授に師事。12年伊東市民病院臨床研修センター副センター長、18年岐阜大医学教育開発研究センター客員教授などを経て、22年より現職。日本医学教育学会認定専門家制度コースワークの実施責任者。同学会学習者評価部会メンバー。



##### ●いとう・しょういち氏

1998年千葉大医学部卒。2003年同大大学院博士課程修了。02年成田赤十字病院、04千葉大助教、09年同大講師などを経て、19年より現職。千葉大において卒前・卒後の医学教育以外にも他職種の研修にも携わり、高等教育センターの副センター長として全学のアセスメントポリシーの作成を行っている。医学教育モデル・コア・カリキュラム令和4年度改訂版において「学修評価」の執筆の取りまとめを担当した。



##### ●にしごり・ひろし氏

1998年名大医学部卒。2008年英ダンディー大医学教育学修士課程、20年蘭マーストリヒト大医療者教育学博士課程を修了。07年東大医学教育国際研究センター、12年京大医学教育推進センターを経て、19年より現職。医学教育モデル・コア・カリキュラム等の次期改訂に向けた調査・研究医学チーム副座長。日本医学教育学会理事補佐。



##### ●さいとう・ゆうご氏

2011年京大教育学部卒。18年同大大学院博士課程修了。博士(教育学)。17年同大高等教育開発推進センター、18年藍野大医療保健学部理学療法学科助教、19年新潟大経営戦略本部教育戦略統括室准教授などを経て、22年より現職。専門は教育学、特に高等教育における教育評価論と教育測定論。



註1: 簡易版臨床能力評価。学生が医療面接や身体診察を行う際に、指導医がその様子を15~20分程度観察評価する。コアカリでは①病歴聴取、②身体診察、③コミュニケーション、④臨床判断、⑤プロフェッショナリズム、⑥マネジメント、⑦総合の観点から尺度評価し、文章によるフィードバックを記載する評価シートが示されている。

註2: Direct observation of procedural skillsの略。学生が臨床手技を行う際に、指導医がその様子を直接観察して評価する。コアカリでは①適応・解剖・手技の理解、②インフォームド・コンセントの取得、③事前の準備、④技能、⑤清潔操作、⑥安全への配慮・援助の要請、⑦手技後の管理、⑧コミュニケーション、⑨プロフェッショナリズム、⑩総合(概略評価)の観点から尺度評価し、記録とフィードバックを行う評価シートが示されている。

註3: パフォーマンスの質を評価するために用いられる評価基準を示した表のこと。評価の観点とレベルを一覧として示しており、レベルごとの評価の基準が説明されている。

#### ●参考文献・URL

- 1) Acad Med. 1990 [PMID: 2400509]
- 2) Med Teach. 2012 [PMID: 22364452]

プレゼンテーションで悩む、すべての医療者・学生へ

## 医療者のスライドデザイン

プレゼンテーションを進化させる、デザインの教科書

研究や発表で使う「スライド」をよりきれいに、よりわかりやすくするための指南書。一般的なデザインのルールはもちろん、医療職者が多様する「数字」「グラフ」「画像」「フローチャート」などに特化した解説も掲載。医療系スライドの多数の実例を示し、具体的な改善方法を提案する。汎用プレゼンテーションソフトで利用できるフォーマットやアイコンのダウンロード、実際の作成過程の動画付き解説などの付録も充実。

小林 啓

医療者のスライドデザイン



輸液・水電解質のリアルに挑め。経験豊富なDr.長澤の思考過程がみえる20症例。

## Dr. 長澤印 輸液・水電解質ドリル

輸液・水電解質のリアルに挑め。経験豊富なDr.長澤の思考プロセスが見える！つまずきやすい輸液や水電解質をDr.長澤が初学者にもわかりやすく解説。1章(総論)で学んだあとは、2,3章(各論)の症例問題を解いて、どんどん実践すべし。わからないところがあったらいつでも1章(総論)に立ち返ろう。解き終えた後は付録の関連検査値・式、逆引き疾患目次、Learning Pointまとめもご活用ください。

長澤 将

