

総説

学習者評価とコンピテンシー基盤型教育

大西 弘高*¹

要旨

現在、医学教育カリキュラムは、コンピテンシー基盤型教育の考え方によって再構成が図られようとしている。コンピテンシーが達成されているかどうかを質管理する際には学習者評価の質が非常に重要だが、どのように質改善を図るべきかについては現状でまとまった資料がない。この論文では、文献的な検討により、学習者評価の観点からコンピテンシー基盤型教育がどうあるべきかに関して論じる。

キーワード：学習者評価，カリキュラム，医学教育

Learner assessment and competency-based education

Hirotaka Onishi *¹

Abstract

Currently, medical educators are endeavoring to reform curricula from the viewpoint of competency-based medical education (CBME). The quality of learner assessment is very important to control quality regarding whether competencies are achieved or not. Presently, however, there are no good articles on how to improve the quality of CBME. In this article, I reviewed references and discuss how CBME should be structured from the viewpoint of learner assessment.

Key Words: learner assessment, curriculum, medical education

現在、わが国の医学教育は分野別外部評価システムをきっかけにして変革が続いている。その中で、コンピテンシー（アウトカム）基盤型教育（competency-based medical education：CBME）は21世紀の新しい考え方として世界医学教育連盟の基準において非常に重要な役割を果たしている^{1, 2)}。

一方、学習者がすべからくコンピテンシーを達成できているかどうかという質管理の観点では、学習者評価の質を高め、改善を図っていく必要がある。しかしながら、医師国家試験や共用試験といった大学を超えた学習者評価の枠組み、各大学で行われるべき様々な学習者評価活動は全体としてCBMEの基盤を構成できているかについて問題が多いと考えられる。

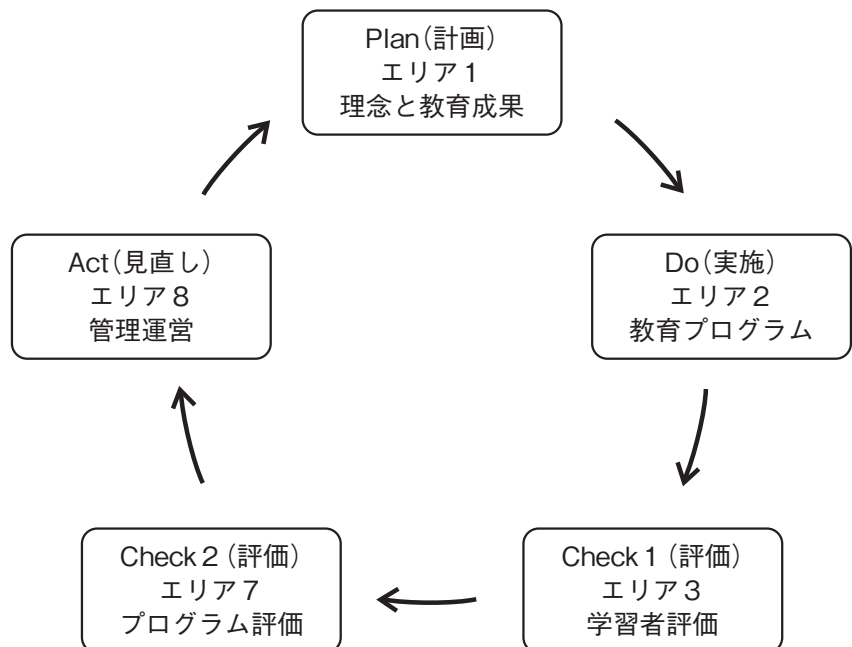


図1 医学教育分野別評価基準日本版とCBMEとPDCAサイクルの関係

*1 東京大学大学院医学系研究科医学教育国際研究センター, International Research Center for Medical Education, Graduate School of Medicine, The University of Tokyo [〒113-0033 東京都文京区本郷7-3-1 医学部総合中央館2F]

受理日：2016年5月8日，採録決定日：2016年6月5日

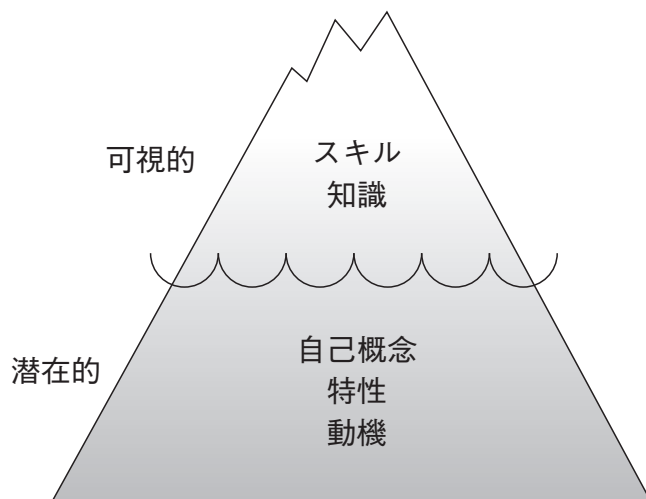


図2 コンピテンシーの冰山モデル

本稿では、CBME の考え方、これに基づいた学習者評価のあり方についてまずは枠組みを明確にし、医学教育領域におけるわが国の現状を踏まえた上で、今後の改善策について提言を行いたい。

1. コンピテンシー基盤型教育(CBME)

医学教育分野別評価基準日本版³⁾によると、エリア1で理念と教育成果（アウトカム）、エリア2で教育プログラム、エリア3で学習者評価、エリア7でプログラム評価、エリア8で管理運営について触れている。これは、教育をシステムに見立て、継続的に計画、実施、評価、見直しといったPDCAサイクルが形成されていると考えられる（図1）。そのとき、Plan（計画）は理念とアウトカム、Do（実施）が教育プログラム、Check（評価）が2段階あって学習者評価とプログラム評価、Act（見直し）が管理運営にそれぞれ該当すると理解することが可能であろう。事実、多くの大学医学部は、何年かに一度何らかのカリキュラム改革を行うものの、それ以外の年には同じカリキュラムを継続的に実施すると共に、小規模な改善を図るという形でカリキュラムを運営している。このような実践が、図1のPDCAサイクルに当たる。

コンピテンシー（アウトカム）基

盤型教育は、教育機関がそのプロダクトであるところの修了者、卒業者のコンピテンシーに説明責任を持ち、質保証するというシステムである。これは、製造業にたとえば、製品の質管理に責任を持ち、製品販売後にも何らかの問題が起これば、リコールによって無償修理したり、PL（product liability）法によって責任をとったりするのと同様である。患者や社会に補償するという考え方にもつながり得る。

ここで、アウトカム、コンピテンシー、コンピテンスといった用語の定義について触れておきたい。White は、コンピテンスを生体がその環境と効果的に交渉する能力ならびにその意欲（有能感や動機づけも含む）と定義した⁴⁾。一方、McClelland は、達成動機の測定方法、職務上の業績を予測できる変数やテスト手法を見出そうとしてコンピテンシーを提唱した⁵⁾。ここでは、コンピテンシーはコンピテンスよりも集合的な概念と位置づけている。さらに、Spencer & Spencer は、コンピテンシーをある職務において卓越した業績を生み出す要因となっている個人の基盤的特徴とし、コンピテンシーには可視的な部分だけでなく潜在的な部分も多いという冰山モデルを提示した（図2）⁶⁾。

一方、Frank らは、これらの用語には混同がみられるとして再定義を

試みた⁷⁾。ここでは、コンピテンシーが明確に評価可能であるという点が重視され、アウトカム基盤型教育ではなくコンピテンシー基盤型教育（CBME）の用語で統一が図られた。これと共に、従来用いられてきたアウトカム領域という言葉に換えてコンピテンス領域という用語が充てられ、コンピテンスを医師の能力全体を表す用語とした。また、コンピテンシーは、コンピテンス領域の下位項目として、細分化した能力であるとされた。上記の White や McClelland の定義と包含関係が逆転している印象もあり、世界的に広く受け入れられているとはいえない部分もあるが、以下この論文ではこの定義に従って記載していく。

さて、医学教育分野別評価への準備をする医学部は、医学教育分野別評価基準のエリア1の基準に従ってCBMEへの第一歩を踏み出すと考えられる。評価基準の「1.4 教育成果（educational outcomes）」には、(a) 基礎医学、(b) 公衆衛生・疫学、行動科学および社会医学、(c) 診療に関連した医療倫理、人権および医療関連法規、(d) 診断手順、手技、コミュニケーション技法、疾病の治療と予防、健康増進、リハビリテーション、臨床推論や問題解決、(e) 生涯学習能力および医師の様々な役割に関連したプロフェッショナルリズム、といった教育成果5項目に関する注釈が述べられている³⁾。各医学部は、最低限これらを含む形で教育アウトカムを設定し、あとは個々の特色を打ち出すことになるだろう。

上述した Frank は、CBME の実施プロセスを、①卒業生に必要な能力（コンピテンス領域）の同定、②コンピテンシーやその要素の明確な定義、③進度に従ったマイルストーン（milestone：一里塚—コンピテンスが成長する際の節目に当たる時期に、どのレベルにいるかを示す明確な記述）の設定、④教育活動、経験、指導方法の選定、⑤マイルストーンを測定する評価手法の選定、⑥アウトカムが達成できたかのプログ

表1 よい学習者評価のクライテリア

(1)妥当性 (validity) : 評価したいものを評価できているか
(2)一貫性 (consistency) : 反復した評価で同様の評価がなされるか
(3)平衡性 (equivalence) : 異なる場や状況で同様の評価がなされるか
(4)実施可能性 (feasibility) : 実際に実施できる可能性が高い
(5)教育効果 (educational effect) : 学習を動機づける効果が高い
(6)触媒効果 (catalytic effect) : 将来の学習につながる効果が高い
(7)受容可能性 (acceptability) : 関係者が評価プロセスや結果を受け入れる

表2 「よい学習者のクライテリア」を改善していく際に注意すべき4つの点

(1)正当性を鑑み、患者や社会の視点を取り込む必要がある
(2)学習者評価、フィードバック、継続学習の相互関係に関する認識を重視したものである必要がある
(3)学習者評価のシステムに向けて発展されていく必要がある
(4)認証評価プロセスに向けて発展されていく必要がある

ラム評価、の6段階に分けている。各医学部は、①については進めることが難しくないと思われる。一方、②～⑥のプロセスについては、例えば現状でも何らかのカリキュラムが実施され、学習者評価が行われていると考えられるため、それらを④や⑤に近づけることと、②や③を優先することのどちらが重要かは難しい問題であろう。

2. 学習者評価のあり方

学習者評価に関する隔年の医学教育学会である Ottawa 会議では、2010年にいくつかのポジションペーパーが出された。よい学習者評価のクライテリアについては、7つの項目を挙げている(表1)⁸⁾。

また、今後「よい学習者のクライテリア」を改善していく際に4つの点に注意すべきであるとしている(表2)。単に学習者の能力を測定、評定するだけでなく、長期的視点で学習、教育の観点を踏まえていること、継続的なシステムや質改善を考えていることが重要と言える。

CBME の考え方を基盤に据えた上で、学習者評価を継続的に改善し

ながら用いられるシステムと位置づけると、図3のような項目の関連性が重要であろう。コンピテンシー領域が設定され、それに従ってブループリント (blueprint) が明示される。

表1で示した一貫性や平衡性、すなわち反復した評価、異なる場や状況での評価において同様の結果が得られることを保証するためには、評価の内容やプロセスの設計図が必要であり、これが学習者評価においてはブループリントと呼ばれている⁹⁾。

作問や実施は、ブループリントに従って行われ(各科が自分たちの考えで好き勝手に実施するのではなく、教育機関としてのガバナンスを持ってコンピテンシーやブループリントに沿って実施する)、採点はMCQのような客観テストでない限り、信頼性検証が行えるように複数採点者によって行い、最終結果は進級判定委員会のような説明責任を負える組織がそのプロセスや結果を保証する。

コンピテンシー領域の中には知識に関わる内容が多いという理由で、MCQを主体的に用いて評価可能な領域もある。一方で、コミュニケーション技法など、実技を中心に評価を行うべき領域もあるし、診療に関

連した医療倫理のように、実際の症例に基づいて、各自の意見を尋ねるような口頭試問が必要な領域もあるだろう。このように、コンピテンシー領域によっては、作問や実施の手間がかなり大きくなる可能性があることに注意が必要である。

これは、例えばMCQのみで実施されている現状の医師国家試験や、それに類した方法では、学習者評価が不十分になってしまうコンピテンシー領域が生じてしまうことを指す。医師国家試験自体は高い信頼性を持って評価すればよいため、特にそのこと自体が大きな問題というわけではない。一方、各教育機関においては自らが設定したコンピテンシー領域の中に、十分に評価できていない領域が残るのであれば、説明責任を果たせていないことになってしまう。

また、医師国家試験においてはブループリント(出題基準)が示されているが、共用試験OSCEにおいては、ブループリントに不明瞭な面が残されている¹⁰⁾。例えば、診察手技とコミュニケーション技法の点数バランスといった点については触れられていないため、各医学部はそれぞれの授業時間をどういう比率で配分すべきかについて方針を立てにくい。

作問、実施、採点においては、各医学部にて領域別に教育し、合否判定を出すべきかどうか悩ましいところだろう。例えば、内科について十分学べているかどうかを判定したいと考えた場合、筆記試験に関しては内科で独立した試験にすることは可能だろうが、コミュニケーション技法やプロフェッショナルリズムといった点については、評価にかけられる人的資源、時間的資源のことを考慮すると、他科と共に評価せざるを得ないと思われる。

合否判定は、卒業や進級の決定という観点からは、厳密さが要求される。単位制によってある科目だけが不合格というような判定が行われることもあるが、関係者が合議によって判定の方が安全であり、学生が



図3 継続的改善を図るための学習者評価システム

ら訴訟を起こされるといった事態にも対応しやすい。その意味では、医学部には単位制の運用自体が馴染まない部分が生じており、学年制を用いた方が説明責任を持った管理・運営がしやすいと思われる。

これらのプロセスは、毎年全体を見渡し、改善点がないか確認すべきである。数年に一度は学習者評価に無事合格した者たちが、長期的にみて本来評価したかったことを満たしているのかどうかを、予測妥当性の見地から確認することも必要である。この作業により、掲げたコンピテンス領域が単なるお題目になってしまうことを防ぎ、実質的な教育の改善につながっていることを示すことができる。

3. 今後の課題

CBMEと学習者評価は表裏一体であり、これらを継続的に改善するためには、医学部全体でかなりのリソースを配分して取り組む必要がある。以下に、筆記試験、OSCE（客観的臨床能力試験）、WBA（現場業務に基づく評価 work-based assessment）、管理・運営上の改善点の4点に分けて述べていきたい。

A) 筆記試験

筆記試験については、いわゆる客観テストとそれ以外の筆記試験がある。客観テストとは、MCQ (multiple choice questions) や EMI (extended matching items) のように自動採点が可能なタイプのテストである。作問は大変だが、正答率や識別指数による分析がやすく、改善を図りやすい。問題としては、浅い学習を惹起しやすいこと、卒業認定や資格試験などの high-stakes test (合否結果が社会的にも重要な意味を持つような試験) においては問題再現を図ろうとする受験者の行動が生じやすいことが挙げられる¹¹⁾。

それ以外の筆記試験は、短所として採点に時間がかかること、採点自体の標準化が難しいことが挙げられ

る。しかし、key feature problem¹²⁾ など臨床推論能力を評価できる手法もあり、ある程度深い問題解決能力を評価できる長所がある。

筆記試験に関してブループリントを設定する際には、まずは客観テストとそれ以外の筆記試験の構成を考えるとところから始まる⁹⁾。受験者は、これらのテストに対して異なる学習方略が適していることを知っていることが多い¹³⁾。よって、どういう形態の評価が行われるかを示すことにより、受験者はそれに合った学習行動をとると予測される。いくつかの教科で構成された試験問題を作るときには、どの教科からどの程度の割合で出題されるのかを示しておくことも必要である。これも、どの領域をどのように学習するか、どの領域により重きを置くかといった受験者の学習行動につながる情報である。

客観テストにおいては、事前のブラッシュアップが必要となる。作問を担当する指導者が実際に解いてみるなど、不適切問題が生じないように十分準備することが推奨される。一方、それ以外の筆記試験においても一定のブラッシュアップは必要だが、それ以上に手早く、正確に採点することの方が重要である。そのため、用語の揺れをどこまで容認するかといった採点基準を明示することが求められる。例えば、全身性エリテマトーデスが正解だった場合、SLEは正解か、全身性狼瘡も正解でよいのか、systemic lupus erythematosus は部分点なのか、など様々な回答パターンを予測して配点を決めておくような工夫が必要である¹⁴⁾。医学教育分野別評価基準の観点では、信頼性検証を行うことも求められる。客観テスト以外の筆記試験においては、一部答案を信頼性検証のために二重で採点し、信頼性検証（評価者間信頼性）することも必要となる (Q3.1.1の基準による)³⁾。

B) OSCE (objective structured clinical examination: 客観的臨床能力試験)

OSCEは、実施に必要なリソースが非常に多くなる。客観性を持たせるためには、一定以上の規模で実施せざるを得ない上に、医学的な専門性を持つ評価者が多数必要になるからである。また、臨床的なシナリオを設定し、標準模擬患者 (standardized patients) を準備する必要も生じる。考慮すべき点が多すぎるため、ともすればどこかで用いられている内容を踏襲するのみという状況も生まれがちである点には注意が必要である。

多くの場合、OSCEで評価できるコンピテンス領域は、コミュニケーション技法と、手技的なスキル（身体診察、侵襲的な治療ないしは検査手技）に分けられる¹⁵⁾。古典的なタキソノミー（学習目標分類学）の考え方からは、知識、スキル、態度の3種類の学習目標のうちスキルや態度を評価すると言われることもあるが、Millerの三角¹⁶⁾（図4）を考慮すると、臨床評価の真正性が高いかがどうか問題である。ちなみにOSCEは上から2番目、Shows howのレベルを評価しているが、Doesのレベルは評価できない (in vitroの臨床評価とも呼ばれる) ので、真の態度に裏打ちされていないと言える。

OSCEのブループリントを作成する際には、まずはコミュニケーション技法と手技的なスキルをどの程度の割合で評価するかという大きな枠組みが重要であろう¹⁷⁾。例えば、共用試験OSCEにおいては身体診察等の技法の方がコミュニケーション技法よりも占める割合が大きい¹⁰⁾。このことから、受験者はコミュニケーション技法よりも身体診察技法の学習に時間を割くといった考え方に至る可能性もある。

作問、実施、採点というプロセスにおいては、課題のどの部分を評価するのか、どのような実技に対して点数が与えられるのかといった点について細かな取り決めが必要である。これは評価者用マニュアルや評価者トレーニングという形で評価者に伝

えられ、当日使用する評価表によって評定につなげられる。ここで、複数の評価者が同時に評価すれば評価者間信頼性の検討が可能となる。

OSCE 全体の構造として内的一貫性があるかという観点では、 α 係数を検討することで確認することが可能である。一般的に、内的一貫性はステーション数に依存していることが知られている¹⁸⁾。ステーション数と評価者数のどちらがどの程度コミットしているかを確認するためには、一般化可能性理論を用いた解析を行うこともある¹⁹⁾。

C) 現場業務に基づく評価 (WBA)

前述した Miller の三角¹⁶⁾ を参照すると、Does, すなわち現場業務に基づく評価 (Work-based assessment: WBA) と呼ばれるような現場での業務を通じた評価が重要であることが浮かび上がる。具体的には、mini-CEX や DOPS (direct observation of procedural skills)、ポートフォリオ評価、臨床実習全般評定尺度が取りざたされることが多い²⁰⁾。

WBA の基本形は、米国の CEX (clinical evaluation exercise)、英国の Long case といった病棟で実際の患者を診察する様子を指導者が評価する方法として開発された。実習の内容に沿っており、正しい学習を動機づけられる点ではよいのだが、1 人の患者を診察する様子を評

価し終えるのに45分かかると言われており、症例数を増やすのが難しいこと、そのために信頼性が低くなることから、総括評価としての位置づけを失っていた²¹⁾。mini-CEX は、外来患者の診療を用いて15分ほどで評価するのが特徴である。10人ほどの診療によって評価できると、信頼性も高いとされている。米国で開発されたが、英国でも公式に用いられるなど、WBA の代表的な評価ツールと位置づけられるようになった。

DOPS は、主に侵襲的な手技に関して、初心者が実際の患者さんを相手に実施する様子を、指導者がチェックリストを持って評価するもので、形成評価、総括評価のいずれにも用いることができる。単に技能の巧拙だけでなく、無菌操作、声かけ、麻酔などの周辺的な内容についても同時に評価を行うことになっている²²⁾。

ポートフォリオは、実際に経験した内容に関しての記述に加え、その内容を再度実施するときによいようにすれば改善できるのかといった振り返りを行って、それらを文章に残していくもので、学習目的に記載した内容が評価にも直接使われる²³⁾。指導者や同僚とのやり取りが振り返りを深めるため、そのような人たちの支援が不可欠でもある。診療録以外にさらに書き記す業務が増えるため、書くことが役立つと思えるところまで十分な支援が必要であるが、

実際にはそのようなレベルにまで達することが簡単ではない。

このような評価手法を加えていっても、コンピテンシ領域の中でプロフェッショナリズムの評価は簡単ではない。一緒に業務をしていれば、業務に対する真摯さ、様々な職種・患者やその家族に接する際の態度といった様々な面が評価できるだろうが、これらの側面に対して、信頼性を持った形で評価できるかという点、一緒にかなりの時間共に働いて初めて可能になるという側面も見えてくる。

これと同様の問題は、臨床実習の評価などでもみられる。全般評定尺度 (global rating scale) を用いた形で、それぞれの科における態度などを中心に評価がなされるが²⁴⁾、誰が評価すればよい評価になるのかという点が大きな問題である。医学生に対しては、他の医学生、研修医、若手指導医などが比較的長時間一緒に業務することが多いだろう。しかし、評価の権限を持つのは教員でしかなく、場合によってはほとんどやり取りをする機会がなかった教授が最終評価を付けることになる。

このように考えると、業務態度を含めた評価は、現場での指導をしているような密に関わりのある指導者が一番よい。その意味で、臨床現場における WBA は、現場指導、フィードバックなどと常に表裏一体であり、評価だけを独立させて議論することは難しいことが理解されるだろう。

逆に言えば、WBA に関しては、筆記試験や OSCE などでは何が評価できず、何を WBA で補完しようとしているのかが重要である。WBA の内容は、現場での業務に大きく依存するため、どのような場でのどのような学習をしているかに立脚し、表 1 の一貫性、平衡性を保証することは困難である。このことから、図 3 は WBA では重要ではなく、図 4 の Miller の三角や、コンピテンシ領域を概観することで、どういう評価が不足しているか、どういう部分を埋

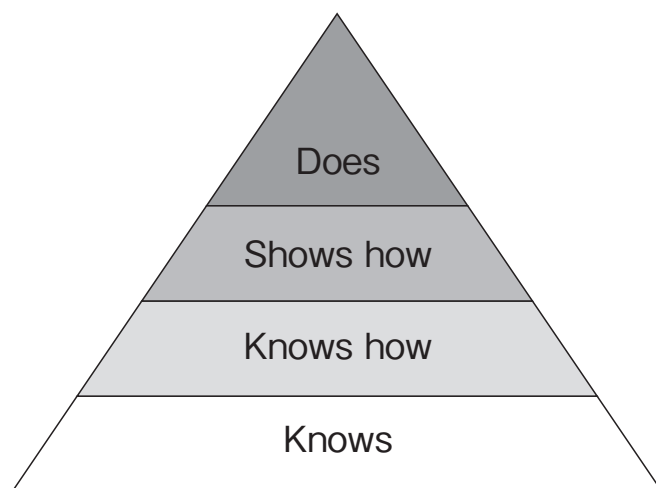


図4 Millerの三角：上の方が臨床評価の真正性が高い

めるべきかを考えるのがよいだろう。

D) 学習者評価の運営・管理

まずは、学習者評価をよいものにしていこうと考えたときに、相応のリソースが必要になるという点に最も配慮が必要である。例えば、OSCE に関しては教員、標準模擬患者、事務担当者などを含め、100人近い人間が半日〜一日拘束されるというイメージで準備が必要である。筆記試験に関しても、コンピテンス領域、ブループリント、作問、実施、採点、最終結果といったそれぞれのプロセスに手間暇をかけると、同様のリソースが必要である。1994年のCusimanoらの研究では、OSCEのコストは学生1人で1ステーションあたり35ドル（試験開発者、事務担当者、標準模擬患者、評価者が全員ボランティアと仮定したら1ドル）とされている²⁵⁾。

WBAに関しては、これを評価だけの取り組みにはできない点に注意が必要である。特に卒前教育においてWBAを導入するためには、実習をかなり診療参加の度合いの高いものにしていくなど、教育自体の改善が不可欠となる。それぞれの科に実習責任者（clerkship director）を配置し、科を跨いだ情報のやり取りを進めるなどの方法がWBAの前提と考えられる。

これらの取り組みには、教育に関する人事や予算といった権限を持つリーダーシップが非常に重要である。多くの医学部では教育担当副学部長といったポジションの人材があてがわれているため、これらの要点に関して理解した上で、必要なリソースを配分することが必要となる。

教育に関する人事や予算に対するリーダーシップにどのようにしてフィードバックを働かせるかは継続的な改善システムを有効に機能させるために必要となる。特に、学生など、教育に関する利害関係の大きいステークホルダーが、圧力を受けずにフィードバックできるようなシステムが構築されているか否かが、肝腎で

あろう。

まとめ

分野別評価のシステムが開始された医学教育領域において、学習者評価がどのようにあるべきかについて、特にコンピテンシーやコンピテンス、教育システムといった観点から論じた。

文献

- 1) Carraccio, C., Wolfsthal, S.D., Englander, R., Ferentz, K., Martin, C. (2002). Shifting paradigms: from Flexner to competencies. *Acad Med*, 77: 361-367.
- 2) Frank, J.R., Mungroo, R., Ahmad, Y., Wang, M., De Rossi, S., Horsley, T. (2010). Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Med Teach*, 32: 631-637.
- 3) 日本医学教育学会分野別認証検討委員会. 医学教育分野別評価基準日本版. 世界医学教育連盟 (WFME) グローバルスタンダード 2012 年版準拠 Ver.1.30. 2015.
- 4) White, R.W. (1959). Motivation reconsidered: the concept of competence. *Psychol Rev*, 66: 297-333.
- 5) McClelland, D.C. (1973). Testing for competence rather than for "intelligence." *Am Psychol*, 28: 1-14.
- 6) Spencer, L.M., Spencer, S.M. (1993). *Competence at work*. New York: Wiley.
- 7) Frank, J.R., Snell, L.S., Ten Cate, O., Holmboe, E.S., Carraccio, C., Swing, S.R., Harris P, Glasgow NJ, Campbell C, Dath D, Harden RM, Iobst W, Long DM, Mungroo, R., Richardson, D.L., Sherbino, J., Silver, I., Taber, S., Talbot, M., Harris, K.A. (2010). Competency-based medical education: theory to practice. *Med Teach*, 32: 638-645.
- 8) Norcini, J., Anderson, B., Bollela, V., Burch, V., Costa, M.J., Duvivier, R., Galbraith, R., Hays, R., Kent, A. (2011). Perrott V, Robert T. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach*, 33: 206-214.
- 9) Hamdy, H. (2006). Blueprinting for the assessment of health care professionals. *Clin Teach*, 3: 175-179.
- 10) 公益社団法人医療系大学間共用試験実施評価機構医学系 OSCE 実施小委員会・事後評価解析小委員会. (2015). 診療参

加型臨床実習に参加する学生に必要とされる技能と態度に関する学習・評価項目（第3.02版）. http://www.cato.umin.jp/06/2016gaku_hyo.pdf (2016年6月4日アクセス)

- 11) Linn, R.L. (2002). Validation of the uses and interpretations of results of state assessment and accountability systems. In: *Large-Scale Assessment Programs for All Students: Development, Implementation, and Analysis*. Eds: Tindal, G., Haladyna, T., Mahwah, N.J.: Lawrence Erlbaum.
- 12) Page, G., Bordage, G., Allen, T. (1995). Developing key-feature problems and examinations to assess clinical decision-making skills. *Acad Med* 70: 194-201.
- 13) Sarnacki, R.E. (1979). An examination of test-wisness in the cognitive test domain. *Rev Educ Res*, 49: 252-279.
- 14) Webber, R.H. (1992). Structured short-answer questions: an alternative examination method. *Med Educ*, 26: 58-62.
- 15) Harden, R.M. (1988). What is an OSCE? *Med Teach*, 10: 19-22.
- 16) Miller, G.E. (1990). The assessment of clinical skills/competence/performance. *Acad Med*, 65: S63-7.
- 17) Mookherjee, S., Chang, A., Boscardin, C.K., Hauer, K.E. (2013). How to develop a competency-based examination blueprint for longitudinal standardized patient clinical skills assessments. *Med Teach*, 35: 883-890.
- 18) Brannick, M.T., Erol-Korkmaz, H.T., Prewett, M. (2011). A systematic review of the reliability of objective structured clinical examination scores. *Med Educ*, 45: 1181-1189.
- 19) Regehr, G., Freeman, R., Hodges, B., Russell, L. (1999). Assessing the generalizability of OSCE measures across content domains. *Acad Med*, 74: 1320-1322.
- 20) Norcini, J.J. (2003). Work based assessment. *BMJ*, 326 (7392): 753-755.
- 21) Norcini, J.J., Blank, L.L., Duffy, F.D., Fortna, G.S. (2003). The mini-CEX: a method for assessing clinical skills. *Ann Intern Med*, 138: 476-481.
- 22) Kogan, J.R., Holmboe, E.S., Hauer, K.E. (2009). Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. *JAMA*, 302: 1316-1326.
- 23) Davis, M.H., Friedman Ben-David, M., Harden, R.M., Howie, P., Ker, J., McGhee, C., Pippard, M.J., Snadden, D. (2001). Portfolio assessment in medical students' final examinations. *Med Teach*, 23: 357-

366.

- 24) Keynan, A., Friedman, M., Benbassat, J. (1987). Reliability of global rating scales in the assessment of clinical competence of medical students. *Med Educ*, 21: 477-481.
- 25) Cusimano, M.D., Cohen, R., Tucker, W., Murnaghan, J., Kodama, R., Reznick, R. A. (1994). comparative analysis of the costs of administration of an OSCE (objective structured clinical examination). *Acad Med*, 69: 571-6.

公 告

医療職の能力開発
編集委員長 大西弘高

前号（第3巻・第2号）にてお知らせしました投稿論文における倫理的な問題の有無に関する検討において、編集委員会側として著者側とやり取りを行って参りました。その結果、著者らは倫理的な懸念事項があったと認め、以下のような謝罪を提示したいという意向でした。

「医療職の能力開発」2巻1号（2013年4月号）に掲載されている「手術チームを対象とした体外循環クロストレーニングの試み—シミュレータ教育における新しい手法—」に関してアンケート調査の説明が不十分ではなかったかとお指摘を受けました。昨今の研究倫理に関する議論を鑑みますと、アンケート調査の際に学術利用や公開に関する説明が十分になされ、公平にご理解を戴くべきであったと反省しております。ご協力くださった皆さま、関係者の皆さまにはご迷惑をおかけいたしましたことを、深くお詫び申し上げます。

この点に関し、調査を受けた側（研究において研究参加者とされていた側）は謝罪を受け入れ、今後このようなことがないように善処を望みたいとのご意見を出されました。

編集委員会としては、以降査読において今回のような研究倫理に悖る問題が生じていないかどうかについても厳しく注意していく所存です。また、研究者の皆さまには、研究開始前、あるいは論文投稿に際し、文部科学省と厚生労働省が平成26年12月に公開した「人を対象とする医学系研究に関する倫理指針」をよくご覧の上、今回のような問題が生じないように対応をお願いします（http://www.lifescience.mext.go.jp/files/pdf/n1443_01.pdf）。