

1. 医学教育における学習者評価

学習者評価は医療者の能力を担保し、医療の質に直結することから、医学教育に関わる者が常に評価活動を検証し、改善を社会に示すことが教育関係者の責務である。COVID-19の影響とテクノロジーの進歩により、新しい評価手法としてシミュレーションの開発や導入が進んでいるが、学習者評価の基盤となる理論、妥当性の概念には大きな変化はない。特に、妥当性は構成概念妥当性で表される単一の概念とされており¹⁻³⁾、学習者評価の質を表す重要な概念である。すなわち、評価を計画・運営するものは、評価すべき能力・行動が測定され評価結果に反映されていることを評価活動の多段階で検討し、評価の及ぼす影響も含めて評価が適切である根拠を示して、妥当性を説明することが求められる。妥当性については、2018年版医学教育白書の学習者評価の記載も参考にいただきたい。

医師のコンピテンシーの修得を目標とするアウトカム基盤型教育がわが国で教育制度に取り入れられ、各教育・研修プログラムに浸透しつつある現在、学習者評価の対象が医師のコンピテンシーになり、さらに実践力を示す entrustable professional activity (EPA)⁴⁻⁶⁾ も評価単位と明記された。

コンピテンシーとは、知識・技能だけではなく、対人関係やプロフェッショナルリズムも含めて、常に、患者や社会のために、様々な状況に対応して、専門職である医師として求められる実践をする能力・姿勢である⁷⁾。したがって、

特定の業務を実践する場面の一回の観察や一つの技能の評価だけで、コンピテンシーを備えていると判定することはできない。特定の業務を専門職として行うために必要な複数のコンピテンシーを発揮し、指導医等の介助や支援を必要とせず実践できるかを異なる状況で繰り返し評価し、その業務の実践力を判定することがEPAである。コンピテンシーとEPAの判定が求められるようになり、評価の対象が知識あるいは技能に限定されていた場合と比較して、計画から判定に至る正しい理解と習熟が必要になった。妥当性に基づいた質の高い評価の実践が求められているのである。

学習者評価には性質の異なる二つの評価が存在する。

- 教育機関が提供する教育・研修プログラムで学ぶ学習者の評価：医学部科目評価、卒業試験、臨床研修プログラム修了認定等
- 医師等専門職の資格認定：医師国家試験、共用試験等

評価の目的、立場の相違により、妥当性の観点から求められる評価の設計、実施方法は大きく異なる。

2. 教育機関が担う学習者評価

(1) 形成的評価の重要性

学習者が修得すべきコンピテンシーは専門職にすべて必須であり、教育プログラムを修了する学習者にはすべての修得を求めるのがアウトカム基盤型教育である。プログラム修了時のコンピテンシー全項目の修得を認定する総括的評価は重要であるが、教育プログラムが機能し、定められた履修期間内に期待されている学習成果をあげるためには、確実な修得を促す形成的評価を履修開始直後より繰り返し実施することが不可欠である。

^{*1} Assessment

^{*2} Masami Tagawa 鹿児島大学大学院医歯学総合研究科医歯学教育開発センター

学習者が評価結果を受け入れ、必要な学習方法を入手し、時間を割いて改善・向上する行動をとるためには、自分の能力を把握して学習の必要性を自覚し、明示された具体的な目標へのモチベーションが高まる必要がある。診療実践の評価（RIME⁸⁾、mini-CEX^{9, 10)}など）、授業中の知識の確認テスト、臨床実習での診療科等実習ユニット終了時のマイルストーン評価、プログレス・テスト¹¹⁾などは形成的評価として有用である。

これらの特徴は、到達すべき目標や最終的に修得していなければならない多様な能力・資質等に対する現在の学習・研修の進捗が客観的に示され、学習者が改善、向上するための有用な情報が提供されることである。評価は合否判定ではなく、点数、ルーブリックの尺度、ナラティブなコメント等で提供される。さらに、有効な評価とするためには、到達目標に対する学習者の現在の位置（学習の進捗度）を絶対的評価基準の尺度上で示し、同一あるいは異なる集団における位置を示す。また、点数や標語では伝わらない学習者の強みや改善点を文章で記述し、面接では対話により伝えることも非常に重要である。忙しい医療の実践を通して学ぶためには、学習者が自信を持ち、困難な状況に対応し、新しいことに挑戦する必要がある、客観的で建設的なナラティブな評価が有効である。

学習者の中には、試験に合格することが、学習目標と考えているものが少なくない。このような学習者が優れた医療者になるためのコンピテンシー、さらには、上のレベルの能力修得を目指して、内的なモチベーションによる好ましい学習を推進するためには、時機を得た形成的評価が欠かせない。また、総括的評価を行う知識、技能の試験以外の学習が見落とされないためにも、すべてのコンピテンシーを形成的に評価し、確実な学習を促す意識は指導者にも求められる。

(2) 臨床実習・研修での評価

学部教育における臨床実習、初期臨床研修、専門医研修は、実践しながらコンピテンシーを

学習する期間であり、形成的評価の対象は、臨床の実践と専門職としての日々の行動である。

また、診断や治療の知識や技能だけではなく、疾患を有した個々の患者の状況に応じた診療の遂行が評価される。ローテーションする診療科の専門領域を学習しつつ、実習・研修期間を通じて一貫した評価を行う。さらに、卒前、卒後の一貫した評価の必要性も示唆されている。

患者の個性、背景、診療の状況を踏まえた個別の実践を観察して、その実践に求められるコンピテンシーすべてが十分に発揮されているかを評価する。学習者自身が何を判断してどの範囲を自分で行い、陪席していた指導者がどの程度の補助を要したのか、事前事後に指導者からどの程度の指示が必要であったのかなどの情報を記録した評価結果は、実践を一人で任せられるかの判断材料となる。実践直後に学習者に評価結果を開示し、優れた実践を褒め、改善を促し、改善方法も討議すると良い。

診療科等での実習・研修の途中および終了時の学習成果の判定では、十分な量の実践の評価データを用いる。信頼性の観点から mini-CEXは何回以上必要かという議論もあるが、ローテーション期間中の能力の向上と、評価した際の診療状況の多様性を考慮すると、むしろ必要な評価の回数を厳密な数字で出すことは好ましいとは言えない。繰り返す評価により新たな情報が得られなくなったときに、十分なデータを得たと考えることも推奨されている²⁾。スコアだけでは判定できないこともあり、診療に参加し、観察していた指導医や同僚の意見を直接聴取して実際の行動を確認し、総合的に判断することが望ましい。臨床実習開始、研修開始、修了、専門研修終了という能力レベル（マイルストーン）を示す一貫した尺度のどの位置にあるかを判断する。

EPA の評価は、実践の習熟度で判断する。記録されたデータから改善が明らかになり、さらに、独立して実践できる次の段階に進められる信頼度が得られたかを総合的に判断する。

(3) プロフェッショナリズムおよびプロフェッショナル・アイデンティティの評価

プロフェッショナリズムには、利他的態度、倫理的対応、患者の尊重、高い専門性の発揮、自学自習など多くが含まれる。プロフェッショナリズムは用語が使用される場面により異なる定義で用いられることがあり、評価にはプロフェッショナリズムを説明したそのプログラムでの明確な定義が必要となる。医学教育モデル・コア・カリキュラムでは、医の倫理と生命倫理、患者中心の視点、医師としての責務と裁量権が示されている。臨床研修の到達目標では、社会的使命と公衆衛生への寄与、利他的な態度、人間性の尊重、自らを高める姿勢と説明され、倫理性は別の資質・能力として記述されている。

いずれにしろ、プロフェッショナリズムもコンピテンシーの一つである。コンピテンシーとは、観察できる行動として記述され、修得と評価が可能と定義づけられている^{12, 13)}ことから、確実な修得を促し保証する評価を行わなければならない。

プロフェッショナリズムの判定にも評価データが必要となる。患者を尊重する態度や倫理的対応は、患者との診療場面の観察、診療録、カンファレンスでの症例提示の場面で明らかになることもあり、看護師等の観察記録によって日常的な対応が評価されることもある。さらに経験の振り返りに、倫理的考え方、姿勢が記述されることで明らかになることもある。日々行動を共にする同僚が、学習習慣や責任ある診療を行うための生活管理を知っているかも知れない。研修当初よりプロフェッショナリズムの評価計画を作成し、観察と記録の機会を確保することで評価が可能になる。

プロフェッショナリズムで求められている行動ができることと、アンプロフェッショナルな行動をすることは別である。アンプロフェッショナルな行動に対する評価・対応は必須であるが、プロフェッショナリズムは評価が困難であるとして、アンプロフェッショナルな行動のみを評価対象とするのは誤解である。

現在では、医師に要求される行動を「自らの

意志で、医師として考え、行動し、感じる人として」示すプロフェッショナル・アイデンティティの獲得というコンピテンシーを超えた人材育成も注目されている¹⁴⁾。



図 改変されたミラーのピラミッド

3. 外部機関等による資格認定の要件と実施

(1) 説明責任を果たす情報開示

医学部教育には、医学教育モデル・コア・カリキュラムという指針はあるが、個々の教育機関が特徴あるプログラムを提供しており、同じものはない。外部機関による評価は、何をいかに学習したかという教育とは切り離して実施されるものであり、実際に教育・研修機関で教育されている、いないにかかわらず、資格認定に必要な能力の有無を判定する。試験機関が評価の方法と合格基準を社会に対して開示する責任を有する。合否判定の社会的責任が大きいことから、評価の質保証、すなわち高い妥当性が求められることは疑いの余地がない。

受験生は、どのような能力がどのような評価方法で評価されるかを十分に理解し、合格のために要求されるレベルを知ることによって受験の準備が可能になり、判定結果を受け入れることができる。患者をはじめとする市民、社会は、どの範囲を評価し能力が保証されているかを知ることによって、その資格を得た医療者を、保証された能力の範囲内で信頼することができる。そのためには、評価対象を「臨床実習で必要とする能力」「医師として必要な知識、技能」という漠然とした説明ではなく、「どのような場面で何がど

の程度できるか」という具体的情報の開示が求められる。

(2) High-stakes examination に求められる

高い妥当性

○内容妥当性、合否判定

共用試験実施評価機構が公表している「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修・評価項目」は、評価の出題範囲を示した資料であり、実際の試験ですべての受験生が全項目を評価されることはない。評価の内容妥当性を保証するために、実際の評価で用いる問題、課題の出題方針を示す試験の設計書、いわゆるブループリントが必要になる。ブループリントに従って評価を実施し、同じ能力基準で合否判定した場合、異なる試験課題の組み合わせで実施された客観的臨床実技試験 OSCE であっても、異なる年度に実施した試験であっても、同等の能力を評価したと説明することができる。同様に、医師国家試験が、毎年同じ能力を評価していると見なすことができる。ブループリントは公表する必要はないが、異なる試験で、同じ能力を担保していることを保証するのは評価機関の役割である。

合否判定による資格の付与を目的とした試験は、スコアの開示はしないことも多い。最低ラインの点数で合格しても満点で合格しても、与えられる資格とそれに伴う義務と責任に違いはない。合否判定基準作成には絶対的に正しい方法があるわけではないとされているが、多くの議論を経て適正であることが認められた方法で慎重に行われることになる。方法の選択と基準作成の手順は、教育学、医学教育の学習者評価の成書を参考にしていきたい。

このように評価の内容妥当性が検証された上で点数、スコアの再現性（信頼性）のデータを示すことで、高い妥当性を有した評価であると説明できるのである。信頼性が高くても内容妥当性が低ければ、その点数あるいは合否判定結果を、能力を示したものとして利用することはできない。

○試験の質を担保するための情報の開示／非開示

試験問題、課題は、実際に評価に使用して初めて問題の特性である難易度（正答率）や能力の識別力が明らかになる。したがって、内容妥当性、信頼性の高い試験を確実に実施するためには、試験問題は実施後公表せず、識別力の高い良問を蓄積して再利用すべきとされている。形成的評価とは全く異なる対応である。現在、医師国家試験問題は実施後公表されているが、これは出題形式や傾向、合格レベルを知る情報源として既出問題が重要であったことを背景に、開示への強い要望が反映された結果であった。既出問題がなくても準備に必要な情報が提供され、実施機関による公式の練習問題が公表されるなど受験生の視点に立って不安を解消することができれば、問題を非開示とすることで high-stakes examination の質保証を進めることも可能である。

OSCE においても、評価方法への習熟度によって評価結果が変わる。すなわち、患者診療の能力ではなくシミュレーション試験への対応能力を評価するのであれば、妥当性は脅かされることになる。例えば、試験室の設定、機材の使用、試験特有の患者の振る舞い、診療の進め方、評価者の陪席、時間配分等、臨床実習等で経験した実際の診療とは異なる状況での対応を求める場合、評価方法への慣れ、習熟で評価点は高くなる。それでも評価のために特殊な対応を受験生に求めざるを得ないのであれば、あらかじめ十分な情報を提供し、評価結果が「慣れ」の影響を受けないようにする。試験の質を担保するために何を開示し、何を非開示とするかは、妥当性の観点から判断されなければならない。

(3) 評価機関の特性

医学教育において、誰がどのように医師の資格を付与し、保証するのかが重要な課題として議論されてきた。わが国では、医師の資格認定の評価機関として、医師免許は厚生労働省、共用試験は医学部歯学部が組織した共用試験実施評価機構、専門医認定は専門医制度に基づいた各専門領域の学会があげられる。国家試験で医

師の資格を認定することで医師に対する社会からの信頼を得て権利を保証してきた。一方でアメリカのように専門職が組織する機関が資格を認定する国もある。これは政治に左右されることなく専門家としての自分達の質を高め、維持する autonomy を尊重した考えに基づいている。COVID-19 の感染拡大で医療崩壊の危機が現実になった時に、医師・医療者の供給と質保証を議論せざるを得なくなった。もし、資格のハードルを下げ、免許保有者数を増加させれば医師・医療者の関わる政策の遂行が容易になるが、それで診療能力を十分に担保できないならば医師・医療者への社会からの信頼を失うことにもなりかねない。共用試験の公的化が決定した現在、様々な学習者評価と試験による評価の限界を理解した上で、医師・医療者育成について誰が責任を持って、どのように資格を認定することが安全で良質な医療の提供を確実に保証し、医師・医療者への社会からの信頼を維持することができるのかを考えることも必要であろう。教育プログラムでの評価と外部機関による評価がそれぞれを補完し、医師・医療者の質保証を行う制度が求められている。

■文 献

- 1) Messick S. Validity. In : Linn RL editor. Educational measurement 3rd ed., American Council on Education, Macmillan, New York, 1989, p.13-104.
- 2) Kane MT. Current concerns in validity theory. *J Educ Measure* 2001 ; **38** : 319-42.
- 3) Downing SM. Validity : on the meaningful interpretation of assessment data. *Med Educ* 2003 ; **37** : 830-7.
- 4) ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ* 2005 ; **39** : 1176-7.
- 5) ten Cate O, Hart D, Ankel F, Busari J, Englander R, Glasgow N, et al. Entrustment decision making in clinical training. *Acad Med* 2016 ; **91** : 191-8.
- 6) Englander R, Flynn T, Call S, Carraccio C, Cleary L, Fulton TB, Garrity MJ, Lieberman SA, Lindeman B, Lypson ML, Minter RM, Rosenfield J, Thomas J, Wilson MC, Aschenbrener CA. Toward defining the foundation of the MD Degree : core entrustable professional activities for entering residency. *Acad Med* 2016 ; **91** : 1352-8.
- 7) Epstein ME, Hundert EM. Defining and assessing professional competence. *JAMA* 2002 ; **287** : 226-23.
- 8) Bloomfield L, Magney A, Segelov E. Reasons to try 'RIME'. *Med Educ* 2007 ; **41** : 1083-111.
- 9) Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise) : a preliminary investigation. *Ann Intern Med* 1995 ; **123** : 795-9.
- 10) Holmboe ES, Huot S, Chung J, Norcini J, Hawkins RE. Construct validity of the miniclinical evaluation exercise (MiniCEX) . *Acad Med* 2003 ; **78** : 826-30.
- 11) Muijtjens AMM, Hoogenboom RJI, Verwijnen GM, van der Vleuten CPM. Relative or absolute standards in assessing medical knowledge using progress tests. *Adv Health Sci Educ* 1998 ; **3** : 81-7.
- 12) McGaghie WC, Miller GE, Sajid AW, Telder TV. Competency-based curriculum development in medical education. An introduction. WHO 1978
- 13) Frank JR, Snell L, Ten Cate O, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education : theory to practice. *Med Teacher* 2010 ; **32** : 638-45.
- 14) Cruess RL, Cruess SR, Snell L, Steinert Y. Amending Miller's pyramid to include professional identity formation. *Acad Med* 2016 ; **91** : 180-5.