

医療系分野における汎用的能力の評価方法の検討

平山 朋子・斎藤 有吾・松下 佳代
(藍野大学) (新潟大学) (京都大学)

A Consideration of the Assessment of Generic Competences in Medical Care

Tomoko Hirayama, Yugo Saito, Kayo Matsushita
(Aino University) (Niigata University) (Kyoto University)

The purpose of this study is to examine the assessment of generic competences in the field of medical care. Universities are now required to make visible their students' learning outcomes at the program level as well as the course level. Specifically, universities should present their students' abilities as goals of their degree programs and verify to what extent the students have acquired them before graduation.

Commercially available standardized tests are often used as an add-on assessment at the program level in many universities. A typical example of those standardized tests is the Progress Report on Generic Skills (PROG). PROG assesses two types of generic competences: "literacy" and "competency." On the other hand, in the field of medical care, the Objective Structured Clinical Examination (OSCE) has been used as an embedded assessment of program-level learning outcomes. OSCE assesses communication with a simulated patient and clinical reasoning, which is a type of problem-solving in the context of medical care.

By analyzing the relationship between the results of these two assessments, we examined whether PROG is able to assess generic competences in the field of medical care. No correlation was found between the "competency" of PROG and that of the communication skills tested by OSCE, which were regarded as similar constructs. Similarly, no correlation was observed between the "literacy" of PROG and clinical reasoning skills, despite their conceptual resemblance. Meanwhile, significant correlations were identified between the "competency" of PROG and the personalities of students as measured by a questionnaire survey. It is reasonable to conclude that it is difficult to assess the generic competences and regard them as program-level learning outcomes in the field of medical care by using PROG.

〔キーワード：学習成果，汎用的能力，PROG，OSCE-R，臨床推論〕

1. 問題背景と研究目的

(1) 問題背景

近年，日本の高等教育においては，プログラムレベルの学習成果の可視化が重視されている。学習成果とは，中央教育審議会の学士課程答申では，「プログラムやコースなど，一定の学習期間終了時に，学習者が知り，

理解し，行い，実演できることを期待される内容を言明したもの」と定義されている（中央教育審議会，2008a）。学習成果の可視化が重視されるようになった背景には，大学のユニバーサル化で多様な学生が入学するようになったことにより，経済・雇用に直結するような能力の要求が高まっていることや，その経済・雇用において大学卒業者に求める能力が高度化したこと，そして，高等教育のグローバル化により学位の国際通用性が求められるようになったこと，などが挙げられる（深堀，2015）。

このような流れを受けて、各大学（あるいは学位プログラム）は、卒業までに身に付けるべき資質・能力について、目標としての学習成果をディプロマ・ポリシー（学位授与の方針）として具体的に示し、それが当該プログラムにおいてどの程度身に付けられたのかを検証すること、すなわちプログラムレベルの学習成果の可視化が求められている。そして、そのような学習成果を設定するための参考指針が日本においてもいくつか作成されている。

例えば、学問分野が多岐にわたる大学教育において、分野を越えた共通性を持つと期待される学習成果として、学士課程答申（中央教育審議会、2008b）では「学士力」が示されている。学士力は、①知識・理解（多文化・異文化に関する知識の理解など）、②汎用的技能（コミュニケーションスキル、数量的リテラシー、情報リテラシーなど）、③態度・志向性（自己管理能力、リーダーシップなど）、④統合的な学習経験と創造的思考力、の4つで構成されている。

一方、各学問分野における目標としての学習成果は、日本学術会議（2010）の「大学教育の分野別質保証のための教育課程編成上の参照基準」（以後、分野別参照基準）において示されている。分野別参照基準は基本的に、(1)当該学問分野の定義、(2)当該学問分野に固有の特性、(3)当該学問分野を学ぶすべての学生が身に付けることを目指すべき基本的な素養、(4)学習方法及び学習成果の評価方法に関する基本的な考え方、(5)市民性の涵養をめぐる専門教育と教養教育との関わり、の5項目によって構成されており、2020年1月現在、32分野で公表されている。このうち、(3)が目標としての学習成果にあたり、①当該分野の学びを通じて獲得すべき基本的な知識と理解、②当該分野の学びを通じて獲得すべき基本的な能力、に大別され、さらに②が、a. 分野に固有の能力、b. ジェネリックスキルに分けて記述されている（日本学術会議、2010、pp.17-20）。

これらの学習成果は、目標であるとともに評価の対象でもある。学習成果の評価はまず、①直接評価と間接評価、②量的評価と質的评价、の2つの軸で整理することができる（松下、2017）。直接評価とは、学習者の知識や能力の表出を通じて、学習成果を直接的に評価する方法であり、標準テストや客観テスト、パフォーマンス評価やポートフォリオ評価が該当する。他方、間接評価とは、学習者による学習行動や学習成果についての自己報告を通じて、学習成果を間接的に評価する方法であり、学生調査やミニッツペーパーなどが該当する。このうち、標準テストや学生調査は量的評価に位置づけられ、統計学的処理がしやすく、主に選抜やアカウンタビリティなどに用いられる。ミニッツペーパーやルーブリッ

クは質的评价に位置づけられ、主に学生の学習指導や改善のための情報を得るのに用いられる。

学習成果の評価にはもう一つの軸がある（松下、2017）。それは機関レベル・プログラムレベルの評価と科目レベルの評価である。プログラムレベル・機関レベルの評価は、大学のディプロマ・ポリシーに対する評価である。では、この各大学のプログラムレベル・機関レベルの評価はどのように行われているのか。2016年度の「学修成果としての調査・測定を行っている事項」に関する調査結果では、回答した大学のうち（ $n=736$ ）、「外部の標準テスト等による学修成果の調査・測定」を使用しているのは35.5%（261校）であった。また、「学生の学修経験等を問うアンケート調査」は23.0%（169校）の大学で実施されていた。一方、「学修評価の観点・基準を定めたルーブリック」による把握は、6.8%（50校）と少ないのが現状である（文部科学省、2019）。このように、日本のプログラムレベル・機関レベルの学習評価では、間接的な量的評価である質問紙調査や、直接的な量的評価である標準テストが多用されていることがわかる。

大学のプログラムレベルの学習成果を可視化する目的で広く実施されている標準テストとして挙げられるのは、PROG（河合塾・リアセック）やGPS-Academic（ベネッセ）である。このタイプの評価は、科目の成績には直接関係しない〈追加型〉の評価として実施される。大学間の比較が可能であることに加え、結果を量的指標によって簡潔に示せるのでアカウンタビリティの要請に応えやすいという強みを持っている。とりわけ多用されているPROGは、専攻・専門にかかわらず大学卒業生として社会で求められる「汎用的能力」を測定する目的で開発されたテストである（PROG白書プロジェクト、2015）。昨今の学習成果の可視化の要請から多くの大学で実施され、2018年7月時点で累計受験者数は66万人にのぼる（PROG白書プロジェクト、2018）。

このような標準テストに対し、多くはないが、直接的な質的评价によるプログラムレベルの評価の実践も報告されている。そのうちの一つとして「重要科目での埋め込み型パフォーマンス評価」（Pivotal Embedded Performance Assessment：PEPA）がある（Matsushita et al., 2018）。PEPAにおける〈重要科目〉とは、その科目の目標がプログラム全体の目標に直結しており、プログラムの中で、学んだ知識やスキルを統合し、高次の能力を育成・発揮することを求める科目を指す。そのような科目における評価は、直接的な質的评价である〈パフォーマンス評価〉との親和性が高い。そして評価基準をプログラムレベルの目標とすり合わせることで、科目レベルの評価に加えて、プログラムレベルの

目標の達成についても同時に調べるという二重の機能を持つ、〈埋め込み型〉の評価である。当該専門分野に固有の能力だけでなく、専門分野の学びを通じて獲得する「汎用的能力」も評価の対象となりえる。

本研究の対象であるX大学医療保健学部理学療法学科では、プログラムレベルの学習成果の評価としてPEPAを実施している。具体的には、医学・薬学・歯学の分野で臨床実習前に行われる共用試験として客観的臨床能力試験に、映像によるリフレクションを加えた理学療法版の客観的臨床能力試験（Objective Structured Clinical Examination Reflection Method：OSCE-R）という評価（平山・松下・西村，2012）を〈埋め込み型〉の評価として実施している。ここでは、卒業後に理学療法士として必要とされる理学療法に固有の能力（臨床技能など）だけでなく、医療専門職として必要な医療系分野の汎用的な能力（コミュニケーション能力や臨床推論能力）も教育目標であり、評価対象である。問題解決や分析的推論といった汎用的なスキルは、学問分野の中で適用されるときこそ、もっともよく学ばれ、評価されるというBanta & Pike（2012）の見解に沿うものである。

このようにPEPAの〈重要科目〉と類似の機能をもつOSCE-Rは、専門分野に根ざしつつ、医療系分野の「汎用的能力」も射程にした評価を行おうとするものである。一方、標準テストであるPROGは専門分野の内容にかかわらない課題によって「汎用的能力」を評価しようとするものである。では、これらの「汎用的能力」の評価で得られた指標にはどのような関係があるのだろうか。結果に食い違いが生じた場合、どう判断すべきなのだろうか。

このような問題意識から本研究ではOSCE-RとPROGとの関連の検討を試みる。ただし、このように変数間の関係性を議論する際、第3の変数による疑似相関関係も考慮する必要がある。大学における学習成果に対して説明力を持つと想定される第3の変数は複数考えられるが、本研究では特にパーソナリティに注目する。パーソナリティの5因子モデル（外向性・調和性・勤勉性・神経症傾向・経験への開放性）と大学の成績との関連は古くから報告されている。例えばメタ分析による研究から、GPAと勤勉性、調和性、経験への開放性が弱い正の相関関係、外向性が弱い負の相関関係を示す（Richardson et al., 2012）。また、特に医療系の学生を対象にした研究では、3年次までの各学年の成績の分散を、パーソナリティによって（説明率は10%に満たないものの）説明可能であり、勤勉性は弱いながらも頑健な正の影響を持つことが見いだされている（Lievens et al., 2002）。このように、学習成果とパーソナリティには一定の関連が見られることが予想される。もし

OSCE-RとPROGが両者ともパーソナリティとの関連をもつのであれば、パーソナリティの影響力を統制することにより、両者の関連をよりクリアに検討することが可能になる。

疑似相関以外の観点からも、パーソナリティとの関連を含めて検討することの意義は主張できる。昨今の資質・能力の議論に関して、パーソナリティは社会的成功を予測する非認知的特性として注目されている（高橋，2016）。しかしその一方で、パーソナリティは「高ければ高いほうがよい」といった一義的な価値を論じることがなじみにくいものである（遠藤，2017）。また、そのような「深く柔らかい部分」を直接的に教育したり価値付けたりする対象とすることへの警鐘を鳴らす研究もある（本田，2005；松下，2010）。特定の評価手法によって得られた学習成果のための指標が、パーソナリティをどの程度反映するものなのかを検討することは、何を教育・評価対象としていくのかの議論の発展のために、意義があると考えられる。

(2) 研究目的

X大学医療保健学部理学療法学科において、理学療法版のOSCE-R、学外施設で実施した臨床実習、PROG、およびパーソナリティの調査を実施した。本研究の目的は、各評価の結果を分析し、これらの評価で得られた指標間にどのような関係があるのかを明らかにすることである。とりわけ、専攻や専門にかかわらず「汎用的能力」を測定できるとされているPROGが、医療系分野の「汎用的能力」を測定できているのかということに焦点を当て、臨床実習成績やパーソナリティとの関連も考慮しながら検討する。

2. 対象と方法

(1) 対象

対象者は、X大学医療保健学部理学療法学科3年生で、臨床実習開始前の2018年9月に、試験にリフレクションとセミナーを組み合わせ実施した「考えるOSCE-R」（以下、OSCE-Rとする）、OSCE-R後に実施した臨床推論の筆記課題、7月に実施したPROGと学習行動調査、11月に実施した臨床実習（4週間）、以上をすべて行った学生（48名）とした。

(2) 理学療法士養成カリキュラムと求められる能力

理学療法士はリハビリテーション医療を担う職業の一つである。理学療法士には、患者の抱える問題点を解決するための知識、技能、態度・倫理観、とりわけコミュニケーション、情報の収集と分析、論理的思考や臨床推論の能力が求められる。

理学療法学科のカリキュラムは、1年次の一般教養科目と基礎科目から始まり、2年次以降は専門科目が増え

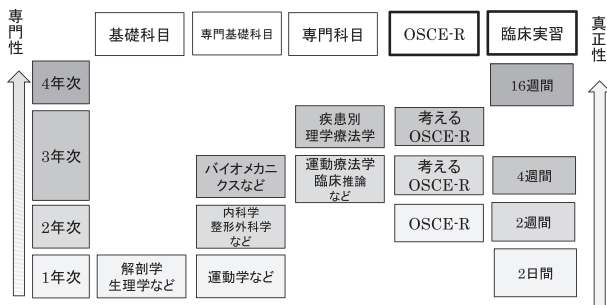


図1 理学療法学カリキュラムと考えるOSCE-R

徐々に専門性が高くなる。またそれに伴いOSCE-Rの課題もより臨床現場の過程に近づけた複雑な設定とし、徐々に真正性が高くなる (図1)。特に各学年に設置している臨床実習は、カリキュラム全体時間数約3300時間のうち3分の1を占める重要な科目である。1年次の臨床実習は、見学を中心に職業理解を深め、2年次の臨床実習は、医療面接や検査測定を部分的に経験し、3年次および4年次の臨床実習は、理学療法評価過程を経験する (図2)。この理学療法評価過程の中に含まれる臨床推論は患者の問題点と解決策を論理的に考える重要な過程である。臨床実習は、患者と関わり、それまでに学習した知識、技能、態度・倫理観を統合しながら学習する機会であり、学生にとっては、その統合した能力を評価される機会でもある。

(3) 分析データ

1) OSCE-R実技試験の評価 (パフォーマンス評価)

X大学理学療法学科のOSCE-Rは、〈第1回OSCE→リフレクション・実技練習 (1週間)→第2回OSCE→リフレクション・セミナー〉というプロセスで構成されている。各試験場には、評価者1名 (教員) とトレーニングされた模擬患者1名を配置し、ビデオカメラ1台を設置して撮影を行う。OSCE-R課題は、筆記課題と実技試験に分かれている。学生は実技試験の60分前に控え教室にて、模擬患者情報をもとに、理学療法評価 (医療面接や検査の進め方など) を考える筆記課題に取り組み、実技試験では、それを25分間で実践し、試験の最後の2分間で実技中の臨床推論について説明を行った。本研究の筆記課題は、3年次前期までに学習した複数科目を統合した整形外科疾患患者の設定で作成した。

OSCE-R実技試験の評価項目は39項目であり、本研究で使用する項目は研究目的に照らし、医療系において汎用性を持つと考えられる「コミュニケーション」「挨拶・導入」「医療面接」の17項目である (表1)。他の項目に関しては、課題・科目固有なものであると判断し、本研究においては使用しない。これら17項目に関して、評価者がチェックリスト方式で評価した結果を用いて、「可」を1点、「否」を0点として得点化した。17項目の

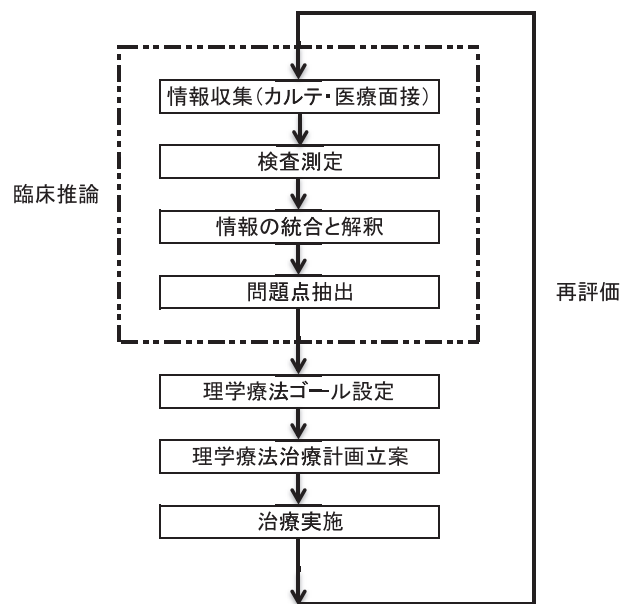


図2 理学療法評価過程

表1 OSCE-R評価表

		評価項目
コミュニケーション	1	身だしなみ・清潔感
	2	適切な顔の向き、目線、表情
	3	丁寧な言葉使い、声の大きさ、話すスピード
	4	話しやすい雰囲気を作ることができる、適切な距離感
	5	共感的な態度
挨拶・導入	6	挨拶及び自己紹介 (大学名・学年・フルネーム) ができる
	7	患者氏名 (フルネーム) の確認ができる
	8	これから実施しようとする内容に関してオリエンテーションが実施できる
	9	これから実施することについての同意を得ることができる
医療面接	10	患者の疾患などの事前情報を元に質問ができる
	11	主訴を聞き出すことができる
	12	症状の変化について質問できる
	13	病前の日常生活活動について質問できる
	14	患者の現在の日常生活活動について質問できる
	15	患者の参加に関する質問ができる
	16	患者の訴えに回答した質問ができる (一方的ではなく対話になっている)
	17	検査測定に必要な情報を収集できる (項目選択と安全性への配慮)
検査測定	18	検査測定の方法は原理原則に従っている
	28	検査測定中にリスク管理ができる (健側から始める、自動運動から始めるなど)
ADLテスト	29	ADLテスト・動作観察を実施する前にリスク管理のための情報収集ができる
	34	動作観察を実施する動作の選択が妥当である
臨床推論	35	医療面接での情報収集について説明できる (統合)
	39	臨床推論に基づいた問題点が概ね説明できる (問題点抽出)

合計得点を算出し、「臨床コミュニケーション能力」とした。

2) 臨床推論の筆記課題 (パフォーマンス評価)

OSCE-R後に臨床推論の筆記課題を実施した (表2)。

表2 臨床推論 (OSCE-R後 筆記課題) (一部抜粋)

【対象者情報 (カルテからの情報)】
 Aさん 男性 または 女性 (模擬患者による) 75歳 両変形性膝関節症 右人工関節置換 (形成) 身長165cm, 体重70Kg
 現病歴: 約15年前に両膝関節の変形性膝関節症の診断を受けている。
 合併症: 高血圧症 (服薬コントロール中), 既往歴: 特になし, レントゲン所見: 関節裂隙の狭小化あり (左側) ステージII K-L grade 3
 ・関節可動域テスト: 右膝屈曲角度: 自動115度, 他動120度, 左膝屈曲角度: 自動100度, 他動120度, 左膝伸展角度: 自動-10度, 他動-5度
 ・徒手筋力テスト
 膝伸展: 左4, 右5 ※筋力テストでは左側のみ痛みあり, 膝屈曲: 左4・右5, 股伸展: 左4・右4, 股外転: 左4・右4, 足関節: 左4 (膝内側に痛みあり)・右5
 ・形態測定: 脚長差なし ・周径差膝蓋骨上縁より10cmと15cm上 1cm右>左 ・FTA: 左側手術前 180° (レントゲンで計測)
 ・痛み 歩行: 左立脚期に特に痛みあり, 特に左膝関節の内側の荷重時の痛みと大腿部前面が突っ張っている感じ, 運動時痛: 左膝関節屈曲の最終→最終域で膝関節につまった感じあり
 ・活動制限について: 歩行200m程度で左膝に痛みが出現する, 右側の靴下が少しはきにくい (両側), 爪が切りにくい,
 ・歩行分析 (独歩)
 歩行は速度が遅く, 歩幅が狭い (左右差はない), 左立脚期に特に痛みあり, 左遊脚相で膝関節が屈曲しにくい, 右遊脚相では右膝関節の側面と大腿部前面が突っ張っている感じあり,
 右側の立脚期に足関節に少し痛みがあり, 立脚期がやや短い, →FTAの変化により, 足部の荷重面の変化が起き, 歩きにくい状態がある。

【筆記課題】 医学的データ追加から臨床推論 (統合と解釈) をし, 問題点, 理学療法ゴールを考える。
 1. 提示された情報や各検査測定データ, 自分の実施した検査測定結果をふまえて, この患者さんの問題点 (心身機能・活動) は何だと考えるか, 重要なデータを提示しながら, 問題点とその問題となり得る理由も説明しなさい。
 2. この患者の問題点を主訴と関連させて考えたとき, 問題点を順位付けし, その理由も述べなさい。
 3. 一般的な症例像の予後を述べ, 列挙した問題点が解決可能かどうかを説明しなさい。

表3 臨床推論 (OSCE-R後 筆記課題) ルーブリック (一部抜粋)

観点	問題解決		論理的思考			文章表現
	背景と問題	主張と結論	根拠と事実	対立意見の検討	全体構成	表現ルール
臨床推論の観点	問題点に至ったデータの読み込み (医学的データ) と理学的検査の意味を読み解き, 問題点を設定する。	問題点としてあげた心身機能・身体構造, 活動制限, 参加制約の順位付けと, その理由を説明できる。	主たる問題点が解決可能かどうか (予後) について根拠を示し, 説明する。	提示した根拠を様々な症例に当てはめたときの「差」について説明, 論証できる。	解決が必要な問題点を明確に示すための過程が論理的に組み立てられ, 表現することができる。	症例レポートの書き方のルールに沿って, 記載できる。
レベル3	心身機能・身体構造, 活動制限, 参加制約の繋がりを説明できる。	心身機能・身体構造, 活動制限, 参加制約における問題点の順位付けとその理由が十分に説明でき, またそれぞれの関連についても説明できる。	一般的な予後を述べ, 提示した問題点が解決可能かどうか十分に説明できる。	教科書の一般症例像と課題の症例の差から特徴的な問題点を見だし, それを十分に説明できる。	抽出した問題点が十分に説明でき, それに対する理学療法プログラムも十分に説明できる。	用紙の使い方が正しく, 誤字脱字もなく, 専門用語を一部正しく使用し, わかりやすい表現で記載されている。引用・参考文献と自分の意見を区別して記載できている。
レベル2	一部省略					
レベル1						
レベル0						

臨床推論能力は, 医療系分野における問題解決能力にあたる。この筆記課題の採点は, 6観点・4水準で作成したルーブリックを用いながら, 教員1名で行った (表3)。なお, 実技試験の理学療法評価過程においても臨床推論が含まれており, 理学療法評価過程の中核をなすものであるが, 実技試験では時間的な制約上, 臨床推論プロセスを詳らかにして評価することは困難である。この筆記課題は, その臨床推論プロセスに焦点を当てたものであり, ルーブリックの4段階であるレベル3~0に依りて得点化した。6観点の合計点を算出し, 「臨床推論能力」とした。

3) 臨床実習の評価 (OSCE-R妥当性検討のための指標)

3年次臨床実習では, 情報の収集と分析, 検査測定, 動作観察・分析を経験し, 患者の問題点を推論し, 問題

の解決策を論理的に考え, 必要とされる理学療法プログラムを立案できるようになることを目標として実施する。臨床実習期間中の指導は施設の理学療法士が臨床実習指導者として行い, 大学教員はメンターとして学生と関わる。臨床実習の成績は, 担当の臨床実習指導者が臨床実習期間の最終週に行った評価や, メンターとの口頭試問, 提出課題などの評価の総合点により算出している。この指標は, 臨床現場, すなわち当学科の学生の大部分にとって将来想定される社会における学生のパフォーマンスを総合的に評価したものであり, 先述したOSCE-Rの妥当性を検討するための指標として適当であると判断した。

4) PROG (標準テスト)

PROGは汎用的能力である「コンピテンシー (周囲の

表4 PROGの構成要素

大分類	中分類		
	要素名	定義	
リテラシーテスト	問題解決力	情報収集力	課題発見・課題解決に必要な情報を見定め、適切な手段を用いて収集・調査、整理する力
		情報分析力	収集した個々の情報を多面的に分析し、現状を正確に把握する力
		課題発見力	現象や事実のなかに隠れている問題点やその要因を発見し、解決すべき課題を設定する力
		構想力	さまざまな条件・制約を考慮して、解決策を吟味・選択し、具体化する力
	言語処理力	語彙や同義語、言葉のかかり受けなど、日本語の適用に関する基礎的な能力	
非言語処理力	数的処理や推論、図の読み取りなど情報を読み解くために必要な（言語以外の）基礎的な能力		
コンピテンシーテスト	対人基礎力	親和性 協働性 統率力	相手の立場に立ち、共感的に接することや、多様な価値観を柔軟に受け入れることができる力。また、チーム内でのお互いの役割を理解し活動ができる力。どんな場面や相手に対しても自分の考えをまとめて伝えることができる力。それが議論の活性化につながることを知っており、周囲にも働きかけられる力
	対自己基礎力	感情制御力 自信創出力 行動持続力	自分の感情や気持ちを状況に合わせてコントロールできる力。また、他者と自分の違いを認め、自分の強みを認識することができ、自分を信頼し、奮い立たせることができる力。一度決めたことや、やり始めたことは粘り強く取り組みやり遂げることができる力。
	対課題基礎力	課題発見力 計画立案力 実践力	さまざまな角度から情報を集め、分析し、本質的な問題を全体的に捉えたり、原因を特定し、課題を抽出することができる力。また、目標設定や課題解決に向けての見通しを立てることができ、計画を見直すことができる力。リスク分析などをしながら実践し、得られた結果に対して検証と改善に取り組むことができる力。

(出典) PROG白書プロジェクト編著 (2018, p.172-173, 付表1) 一部修正。

環境と良い関係を築く力)」と「リテラシー（自分を取り巻く環境に実践的に対処する力）」を測定する目的で開発されたテストである(表4)。どちらも複数の下位構成概念に分けられ、テストへの回答をもとに、潜在ランク理論によって特定の段階のランクで評価される。下位構成概念をすべて含めると変数が20近くと煩雑になるため、本研究ではPROGの特徴を損なわず、かつある程度細かく分析するため、「コンピテンシー」(7段階)及びその下位構成概念の「対課題基礎力」「対人基礎力」「対自己基礎力」(7段階)、「リテラシー」(7段階)及びその下位構成概念のうち問題を解決する力に該当する「情報収集力」「情報分析力」「課題発見力」「構想力」(5段階)の9変数を使用する。

5) パーソナリティ (外的な指標)

パーソナリティの測定は、日本語版 Ten Item Personality Inventory (TIPI-J:小塩他, 2012)を使用した。外向性・調和性・勤勉性(誠実性)・神経症傾向・経験への開放性のビッグファイブ・パーソナリティを各2項目で測定する尺度である。調査対象者にインターネットを介して(希望者には調査用紙を配布して)回答を求めた(「まったく違うと思う(1点)」~「強くそう思う(6点)」の6件法)。各2項目の平均値を求め、尺度得点とした。

6) 変数の対応と分析

まず、上述の手続きによって得られたOSCE-Rの「臨床コミュニケーション能力」「臨床推論能力」の信頼性及び妥当性(「臨床実習成績」との関連)を検討する。次に、OSCE-RとPROGの「コンピテンシー」「リテラシー」及びそれらの下位構成概念、パーソナリティの各指標に関して相関分析を行う。PROGはサンプル問題以

外が非公開のため、厳密な対応付けは困難であるが、それぞれの構成概念の定義を考慮すれば、OSCE-R「臨床コミュニケーション能力」がPROG「コンピテンシー」及びその下位構成概念と、OSCE-R「臨床推論能力」がPROG「リテラシー」及びその下位構成概念と、一定の対応関係にあるものと想定し、関連の検討と考察を行う。

3. 結果

(1) 基礎統計量とOSCE-Rの信頼性と妥当性

本研究で使用した全変数の基礎統計量を表5、OSCE-Rの信頼性と臨床実習成績との相関係数を表6に示す。表6より、OSCE-Rの「臨床コミュニケーション能力」「臨床推論能力」はともに信頼性係数が0.7以上、かつ「臨床実習成績」と小~中程度の有意な正の

表5 基礎統計量

変数名		平均値	SD	最小値	最大値
OSCE-R	臨床コミュニケーション	11.58	2.97	3.00	16.00
	臨床推論	9.44	2.84	2.00	17.00
	臨床実習成績	84.88	6.14	60.00	97.00
PROG	コンピテンシー	2.86	1.67	1.00	7.00
	対人基礎力	3.00	1.58	1.00	7.00
	対自己基礎力	3.14	1.73	1.00	7.00
	対課題基礎力	3.28	1.53	1.00	7.00
	リテラシー	4.28	1.59	1.00	7.00
	情報収集力	3.00	1.36	1.00	5.00
	情報分析力	3.32	1.19	1.00	5.00
	課題発見力 構想力	2.94 3.76	1.46 1.30	1.00 1.00	5.00 5.00
パーソナリティ	外向性	3.33	1.16	1.00	6.00
	協調性	2.97	0.82	1.00	6.00
	勤勉性	2.94	1.15	1.00	5.00
	神経症傾向	3.96	1.05	1.50	6.00
	経験への開放性	3.15	1.13	1.50	6.00

表6 OSCE-Rの信頼性と臨床実習成績との相関係数

変数名	信頼性係数 (α 係数)	臨床実習成績 との相関係数
OSCE-R	臨床コミュニケーション	.73
	臨床推論	.79
PROG	コンピテンシー	-.07
	対人基礎力	-.08
	対自己基礎力	-.09
	対課題基礎力	.12
	リテラシー	.18
	情報収集力	.17
	情報分析力	.09
	課題発見力	.17
	構想力	.22

* $p < .05$

表7 OSCE-RとPROGとの相関係数及びパーソナリティを統制した偏相関係数

	相関係数		偏相関係数 (統制変数: パーソナリティ)		
	OSCE-R		OSCE-R		
	臨床コミュニケーション	臨床推論	臨床コミュニケーション	臨床推論	
PROG	コンピテンシー	.05	.03	-.15	-.07
	対人基礎力	.10	-.03	-.05	-.10
	対自己基礎力	-.09	.03	-.31*	-.07
	対課題基礎力	-.03	.13	-.10	.18
	リテラシー	.12	.07	.07	.02
	情報収集力	.14	.12	.13	.11
	情報分析力	-.03	.06	.00	.08
	課題発見力	-.08	.08	-.12	-.03
	構想力	.00	.08	-.07	.13

* $p < .05$

相関関係を示しており、信頼性・妥当性が一定程度担保されていると判断できる。

PROGの信頼性・妥当性の検討に関しては、PROG白書プロジェクト(2015)を参照されたい。なお、PROGの諸変数と「臨床実習成績」の間に有意な相関関係は見いだせなかった。

(2) OSCE-RとPROGの関連、およびパーソナリティとの関連

次に、OSCE-RとPROGとの相関係数及びパーソナリティを統制した偏相関係数を表7に、パーソナリティとOSCE-R及びPROGの相関係数を表8に示す。表7から、それぞれに有意な相関関係は見いだせず、ほぼ無相関であった。またパーソナリティを統制した偏相関係数についてもOSCE-RとPROGとはほぼ無相関である。唯一、OSCE-R「臨床コミュニケーション能力」とPROG「対自己基礎力」に有意な弱い負の相関関係が見られた。

表7から、OSCE-Rとパーソナリティはほぼ無相関～弱い相関である。それに対して、PROGの「コンピテンシー」とその下位構成概念「対人基礎力」「対自己基礎力」は、パーソナリティの「外向性」「勤勉性」「神

表8 パーソナリティとOSCE-R及びPROGの相関係数

		外向性	協調性	勤勉性	神経症傾向	経験への開放性
OSCE-R	臨床コミュニケーション	-.08	-.23	.03	-.20	.25
	臨床推論	.23	-.05	.19	.00	.03
PROG	コンピテンシー	.62*	.13	.49*	-.31*	.61*
	対人基礎力	.63*	.11	.40*	-.21	.60*
	対自己基礎力	.60*	.25	.39*	-.43*	.62*
	対課題基礎力	.17	.09	.37*	.00	.27
	リテラシー	.21	-.26	.05	.14	.23
	情報収集力	.12	.04	.13	-.05	.18
	情報分析力	-.02	-.08	-.11	.13	.08
	課題発見力	.01	-.31*	-.01	.04	.05
	構想力	.10	-.31*	.13	.36*	.10

* $p < .05$

表9 重回帰分析結果

独立変数	コンピテンシー 対人基礎力 対自己基礎力			VIF
	標準偏回帰係数			
外向性	.36*	.42*	.31*	1.33
協調性	.03	.06	.12	1.41
勤勉性	.25*	.13	.12	1.15
神経症傾向	-.08	.04	-.18	1.51
経験への開放性	.35*	.38*	.40*	1.46
$F(5, 42)$	11.63*	9.55*	11.55*	
R^2	.58	.53	.58	
調整済み R^2	.53	.48	.53	

* $p < .05$

経症傾向」「経験への開放性」と中～強い相関が認められた。このため、統制する前の相関係数と異なる傾向が一部の組み合わせに見られたと考えられるが、概ね、OSCE-RとPROGに有意な関連は見いだせなかったと判断できる。

(3) パーソナリティによってどの程度説明されるか

最後に、OSCE-RとPROGがパーソナリティによってどの程度説明されるのかを、パーソナリティを独立変数とした重回帰分析によって検討した。ただし、パーソナリティと強い相関が複数見られた「コンピテンシー」及びその下位構成概念である「対人基礎力」「対自己基礎力」以外を従属変数にした場合、パーソナリティからの説明力は弱く(調整済み R^2 が最大で.12)、モデルも有意ではない。よってパーソナリティからの説明力が強く、モデルも有意であった「コンピテンシー」「対人基礎力」「対自己基礎力」を従属変数とした重回帰分析結果をまとめたものを表9に示す。

表9から、 R^2 値及び調整済み R^2 値が0.5～0.6程度であった。これは、従属変数に対するパーソナリティの分散説明率が50%程度あることを示し、言い換えれば、「コンピテンシー」「対人基礎力」「対自己基礎力」はパーソナリティによって半分近く説明される指標であるということになる。また、パーソナリティのうち特に「外向性」「経験への開放性」がいずれにも有意な正の標準偏回帰係数を示した。また、「コンピテンシー」はそ

れに加え、「勤勉性」も同様の標準偏回帰係数を示した。すなわち、PROGの「コンピテンシー」「対人基礎力」「対自己基礎力」は、パーソナリティの「外向性」または「経験への開放性」が高い受験者ほど、ランクが高くなる傾向にある。さらに「コンピテンシー」は「勤勉性」が高い受験者ほど、ランクが高くなる傾向にある。

4. 考察

本研究の分析では、理学療法分野の学習を通して習得が期待される医療分野の汎用的能力を評価するOSCE-Rと、専門分野にかかわらない汎用的能力を評価するPROGとの間には、ほとんど相関が見られなかった。OSCE-RはX大学理学療法学科におけるディプロマ・ポリシーに沿うプログラムレベルの学習成果を捉えようとするものであり、臨床現場、すなわち当学科の学生の大部分が将来関与すると想定される状況における学生のパフォーマンスを総合的に判断する臨床実習成績との関連から見ても、その妥当性は主張できよう。つまり、PROGは、少なくとも当学科で涵養を目指す医療分野の汎用的能力を測定していないという疑義が生じる。もちろん、OSCE-RとPROGは、それぞれ測定しようとする構成概念が厳密に一致するように作成されているわけではないため、お互いに強い説明力を持っていなくても不自然ではない。しかし、PROGが専門にかかわらず社会で必要な「汎用的能力」を測定しているとするならば、医療分野における「汎用的能力」と一定程度の関連が見られて然るべきである。では、なぜほとんど相関が見られなかったのか、PROGのサンプル問題と医療系分野で求められるリテラシーとコンピテンシーについて、OSCE-R、臨床実習での課題や評価、カリキュラムや理学療法士の職業特性なども比較しながら検討する。

(1) PROG (リテラシー) と医療系分野で必要な問題解決能力

PROGの「リテラシー」の問題解決能力にかかわる問題は、「情報収集力」「情報分析力」「課題発見力」「構成力」の4つの視点で作成されている。4つのうちの「課題発見力」のサンプル問題は、ある洋菓子屋の売上低迷原因についてロジックツリーを使って探求する問題が紹介されている。これはロジックツリーの空白部分に、5つの選択肢（例えば、広報・宣伝活動の不足）から、原因を当てはめる問題である（PROG白書プロジェクト、2015）。このように、「リテラシー」は専門分野にかかわらない、問題解決のための認知的スキルを評価するため、専門分野の内容や文脈には依存しない問題によって構成されている。今回の分析結果にもとづけば、このような問題で構成されたテストの得点の高さは、医療系分野の臨床現場で求められるような臨床推論（問題解決の

ための思考）の能力を保証していないといえる。

医療系分野における臨床推論（問題解決のための思考）は、患者の病気が同じであっても、年齢、性別、検査データ、患者の社会的背景など、あらゆる情報を統合し、問題の発見と解決方法を考える高次の認知的能力である。すなわち、OSCE-Rで評価対象にした「汎用的能力」は、医療系分野の専門的な文脈をふまえたうえで発揮されると考えられる。「リテラシー」のように専門分野にかかわらない認知的スキルのみでは、臨床推論のような医療系分野の高次の認知的能力を説明しきれないため、ほとんど関連が見られなかったと思われる。

(2) PROG (コンピテンシー) と医療系分野で必要なコミュニケーション能力

PROGの「コンピテンシー」のサンプル問題は、A：初対面の人と話すときでも相手と距離をおかず親しく接する、B：初対面の人と話すときには距離をとって礼儀正しく接する、のどちらにより当てはまるかを選択させる両側選択形式の問題が紹介されている。この問題に関して、社会で活躍する若手ビジネスパーソン（4000人）とランダムに抽出した大学生（1000人）の回答を比較したところ、若手ビジネスパーソンはAを選択する人が多い傾向にあり、大学生との間に有意差があったという。このようにPROGの「コンピテンシー」の正解は、若手ビジネスパーソンの回答比率が高い方に設定されている（PROG白書プロジェクト、2015）。このように設定された正解によって得点化した指標は、「社会で活躍する若手ビジネスパーソンとどの程度傾向が似ているか」を変数化するものともいえる。

しかしながら、理学療法士養成場面で考えると、病気や怪我などでどのような心境になっているのかわからない患者に対して、初対面の段階で距離をおかずに親しく接するように教えることはない。治療を円滑に進めるためにも、まずは患者に不快感や不信感を与えないように礼儀正しく、慎重に、十分に配慮して接するように教える。あくまでもサンプル問題から推論した限りではあるが、PROGの「コンピテンシー」は社会で活躍している若手ビジネスパーソンと大学生を対象にした調査から正解を設定しており、それが医療系分野における正解としても設定できるわけではないということである。

『PROG白書2018』（PROG白書プロジェクト、2018）では、医療・福祉分野は他職種よりも「コンピテンシー」が低いという分析がなされている。しかしながら、サンプル問題のように、想定されている正解が医療系分野内において了解されている正解とは逆となるような設問を中心に「コンピテンシー」が構成されているとすれば、医療系分野で他職種より「コンピテンシー」が低くなってしまふことは十分にありうる。

このようにPROGの「コンピテンシー」は、汎用的と謳われていたとしても、あくまで基準は若手ビジネスパーソンにある。今回の結果は、それを医療系分野で必要とされるコミュニケーションに読み替えることの危険性と、PROGの「汎用性」がどこまでの範囲を意味するのか慎重に見極める必要性を提示している。

(3) PROGとパーソナリティの関係

PROGの「コンピテンシー」に関しては、正解の設定の仕方以外にも議論すべきことがある。「コンピテンシー」は、パーソナリティによって半分程度の分散が説明可能であった。とりわけ外向性、勤勉性、経験への開放性との関連が強い。特に学部を限定せず、大学入学時のジェネリックスキルの規定する因子について調査した研究(亀野, 2016)では、友人との交流が多い学生ほど、PROGの「コンピテンシー」が高く、また「何ごとにも積極的に取り組んでいる」「クラスやゼミ、課外活動などで自発的に行動している」学生ほど「コンピテンシー」が高かったと報告している。本研究におけるPROGの「コンピテンシー」が外向性、勤勉性、経験への開放性で説明できるという結果は、こうした先行研究の結果と一致する。つまり、今回見られたパーソナリティとの関係は、対象が医療系分野であることを原因としているわけではないと考えられる。

このように、PROGではパーソナリティによって多くが説明される指標によって、汎用的能力のランク付けが行われていることになる。プログラムレベルの学習成果の可視化の指標として「コンピテンシー」を用いるということは、当該プログラムにどのような意味を持つのか。また、本田(2005)が指摘するような、学生の「深く柔らかい部分」を価値付けている可能性はないのか、注意深く検討する必要がある。

(4) 分野固有性と汎用性

本研究で得られた結果によれば、PROGは、リテラシー、コンピテンシーともに、医療系分野のプログラムレベルにおける評価とはほとんど相関がなかった。ここで、「医療系大学だから特別である」という反論がなされるかもしれない。しかし、PROGが専攻や専門性にかかわらず、すべての大学生の汎用的能力を測定できることを謳っているのであれば、医療系分野の汎用的能力も測定できるはずである。だが本研究では、PROGではプログラムレベルの学習成果は把握できないことが示唆された。

では、PROG、OSCE-R・臨床推論の汎用的能力は、どのような特徴が見られるのであろうか。松下(2019)によると、汎用性は、①分野固有性に依らない汎用性、②分野固有性を捨象した汎用性、③分野固有性に根ざした汎用性、④メタ分野的な汎用性、の4つに分類され

る。PROGは、その開発者からは、分野を越えた幅広い応用可能性としての汎用性とされており、だとすれば、「①分野固有性に依らない汎用性」に分類されることになる。しかし、少なくとも医療系分野の汎用的能力は測定できていないという特徴を考えれば、「②分野固有性を捨象した汎用性」に分類されるだろう。

一方、学内で実施するOSCE-Rや臨床推論では、医療系分野に必要なコミュニケーション能力や、情報の収集と分析、問題の発見と問題解決の提案といった問題解決能力に関するプログラムレベルの評価と育成がなされており、これらを学生が学外の臨床実習での活動に適用・拡張させていくという過程は、「③分野固有性に根ざした汎用性」を目指すものだといえるであろう。学生は、OSCEの様子を映像で確認しながら、医療専門職としてのコミュニケーションのあり方や臨床推論について振り返る。患者とのコミュニケーションについては、ロールプレイなどの方法で繰り返し練習し、実際の臨床実習の現場に適用させていく。また、臨床実習では病院スタッフと専門用語を使用しながら議論をすることもあり、学生にとっては様々な人とコミュニケーションを図る機会となる。その中で、実際の患者とのコミュニケーションは、OSCE-Rにおける模擬患者とのコミュニケーションとは異なる面もあることに気づいていく。さらにこれらは、学生が自分の日常生活の中で、挨拶や、他者との円滑なコミュニケーションのとり方について見直すことにもつながっていることが報告されている(平山他, 2012)。

このようにOSCE-Rや臨床推論は、学内における模擬場面ではあるが、臨床実習の実践につながり、さらには自分の日常の言動や学習への取り組みの見直しも促すなど、特定の場面に限定せず適用・拡張することが可能であり、その意味で「分野固有性に根ざした汎用性」を有している。今回の調査では、医療系分野で育成された汎用的能力が医療系分野を越えた場面にまで適用・拡張する可能性があるかどうかは確認できなかったが、PROGで高得点を取ることができる汎用的能力は、少なくとも医療系分野の教育目標とは齟齬をきたすおそれがあることが明らかになった。

5. 結論

本研究により、プログラムレベルの学習成果の可視化のために実施されている標準テストのPROGによって、医療系分野における汎用的能力を測定することは困難であることが示唆された。プログラムレベル・機関レベルの評価においては、PEPAのように重点科目におけるパフォーマンス評価も組み込みながら、当該分野の専門的能力と大学卒業時までには習得すべき汎用的能力を評価す

ることが望ましいのではないかと考えられる。しかし、専門分野の学習を通してどのように汎用的能力が育成されるのか、また、その汎用的能力は別の分野へ適用・拡張が可能であるのか、そして、その能力はどのように評価できるのかについては、本研究では十分検討できなかった。今後の検討課題としたい。

謝辞

本研究において、調査に協力して下さったX大学の教員の皆さま、学生の皆さまに感謝を申し上げます。

また本研究は、JSPS科研費（課題番号19K14275）の助成を受けたものです。

文献

- Banta, T., & Pike, G. (2012). Making the case against - One more time. In R. Benjamin et al., *The seven red herrings about standardized assessments in higher education* (pp. 24-30). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
- 中央教育審議会 (2008a) 『学士課程教育の構築に向けて (答申) 用語解説』 (https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1217067.htm) (2020年1月20日)
- 中央教育審議会 (2008b) 『学士課程教育の構築に向けて (答申)』
- 遠藤利彦 (2017) 『非認知的 (社会情緒的) 能力の発達と科学的検討手法についての研究に関する報告書』, 国立教育政策研究所.
- 深堀聰子 (2015) 「アウトカム重視の大学教育改革—その背景と概念の整理」 深堀聰子編著 『アウトカムに基づく大学教育の質保証—チューニングとアセスメントにみる世界の動向』 東信堂, pp.3-32.
- 平山朋子・松下佳代・西村敦 (2012) 「理学療法学を主体的に学ぶ—『OSCEリフレクション法』の試み—」 小田隆治・杉原真晃 (編) 『学生主体型授業の冒険2 予測困難な時代に挑む大学教育』 ナカニシヤ出版, pp.202-221.
- 本田由紀 (2005) 『多元化する「能力」と日本社会—ハイパー・メリトクラシー化のなかで』 NTT出版.
- 亀野淳 (2016) 「大学入学時のジェネリック・スキルを規定する要因分析—北海道大学1年生に対する調査結果をもとに—」 『高等教育ジャーナル—高等教育と障害教育—』 23, 71-78.
- Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students' personality characteristics and academic performance: A five-factor model perspective. *Medical education*, 36(11), 1050-1056.
- 松下佳代 (2010) 「序章 〈新しい能力〉 概念と教育—その背景と系譜」 松下佳代編著 『〈新しい能力〉 は教育を変えるか—学力・リテラシー・コンピテンシー』 ミネルヴァ書房, pp.1-42.
- 松下佳代 (2017) 「学習成果とその可視化」 『高等教育研究』 第20集, 93-112.
- Matsushita, K., Ono, K., & Saito, Y. (2018). Combining course- and program-level outcomes assessments through embedded performance assessments at key courses: A proposal based on the experience from a Japanese dental education program. *Tuning Journal for Higher Education*, 6 (1), 111-142.
- 松下佳代 (2019) (印刷中) 「汎用的能力を再考する—汎用性の4つのタイプとミネルヴァ・モデル—」 『京都大学高等教育研究』 第25号.
- 文部科学省 (2019) 『平成28年度の大学における教育内容等の改革状況について (概要)』 (https://www.mext.go.jp/a_menu/koutou/daigaku/04052801/_icsFiles/afieldfile/2019/05/28/1417336_001.pdf) (2020年1月20日)
- 日本学術会議 (2010) 『回答：大学教育の分野別質保証の在り方について』 (<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-21-k100-1.pdf>) (2020年1月20日)
- 小塩真司・阿部晋吾・カトローニ, ピノ (2012) 「日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み」 『パーソナリティ研究』 21(1), 40-52.
- PROG白書プロジェクト編著 (2015) 『PROG白書2015 大学生10万人のジェネリックスキルを初公開』 学事出版株式会社.
- PROG白書プロジェクト編著 (2018) 『PROG白書2018 企業が採用した学生の基礎力とPROG研究論文集』 学事出版株式会社.
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological bulletin*, 138, 353-387.
- 高橋雄介 (2016) 「パーソナリティ特性研究をはじめとする個人差研究の動向と今後の展望・課題」 『教育心理学年報』 55, 38-56.